ÁLGEBRA LINEAL

PARA ESTUDIANTES DE INGENIERÍA Y CIENCIAS

ÁLGEBRA LINEAL

PARA ESTUDIANTES DE INGENIERÍA Y CIENCIAS

ÁLGEBRA LINEAL
PARA ESTUDIANTES DE INGENIERÍA Y CIENCIAS

Juan Carlos Del Valle Sotelo
Instituto Tecnológico y de Estudios Superiores
de Monterrey, Campus Estado de México

MÉXICO • BOGOTÁ • BUENOS AIRES • CARACAS • GUATEMALA • MADRID • NUEVA YORK
SAN JUAN • SANTIAGO • SÃO PAULO • AUCKLAND • LONDRES • MILÁN • MONTREAL
NUEVA DELHI • SAN FRANCISCO • SINGAPUR • ST. LOUIS • SIDNEY • TORONTO

Director General México: Miguel Ángel Toledo Castellanos
Editor sponsor: Pablo Eduardo Roig Vázquez
Coordinadora editorial: Marcela I. Rocha Martínez
Supervisor de producción: Zeferino García García

ÁLGEBRA LINEAL PARA ESTUDIANTES DE INGENIERÍA Y CIENCIAS

Edificio Punta Santa Fe
Prolongación Paseo de la Reforma 1015, Torre A
Piso 17, Colonia Desarrollo Santa Fe,
Delegación Álvaro Obregón
C.P. 01376, México, D. F.
Miembro de la Cámara Nacional de la Industria Editorial Mexicana, Reg. Núm. 736

ISBN: 978-970-10-6885-4

1234567890 1098765432101

Impreso en México *Printed in Mexico*

The McGraw-Hill Companies

A la memoria de Esther, mi amada madre;
a mi hermano Manuel;
a mis hijas Miriam y Samantha

En un universo quizá infinito
inconcebiblemente antiguo
es una dicha saber que tengo mi origen
en una amorosa madre y en un hermano
que me cuidó como a un hijo
y por eso es mi padre
y percibir una infinitésima parte de mí
en la mirada de dos pequeños seres
que en momentos difíciles
han sido tan grandes.

Contenido

PARTE II
ESPACIOS VECTORIALES, PRODUCTO INTERIOR, NORMAS, VALORES Y VECTORES PROPIOS

PARTE III
APLICACIONES, USO DE TECNOLOGÍA, MÉTODOS NUMÉRICOS

Agradecimientos

Deseo primeramente agradecer a Miguel Ángel Toledo y a Ramón Orduña, quienes me invitaron a realizar este proyecto con McGraw-Hill, por el gran apoyo y paciencia que tuvieron desde el inicio hasta la culminación de la obra, sin ellos hubiera sido imposible terminarla. Este libro lo escribí en el procesador de texto matemático y científico LaTeX y el trabajo editorial para su formación fue considerable; deseo dar las gracias a Marcela Rocha y a Pablo Roig por todo el esfuerzo que hicieron para que el proyecto pudiera editarse en ese formato y por toda la ayuda que me brindaron en el transcurso de su elaboración.

La mayoría de la las figuras las construí utilizando los programas LaTeXPiX, TeXCad, GNUPLOT, TeXCad32 o LaTeX-CAD; deseo dar crédito y reconocimiento a los autores de estos paquetes —de distribución gratuita— por la magnífica tarea que han realizado en esas herramientas de dibujo en el ambiente LaTeX, las cuales facilitaron enormemente el trabajo gráfico en este libro. También quiero reconocer la excelente labor de maquetación por parte de Mercè Aicart Martínez.

Las imágenes 3D —la máquina de la página 416 y los depósitos interconectados de la figura 6-20—, fueron diseñadas por Ernesto Byas Lizardo y Ramón Nuñez Serrania. Todos los dibujos de los circuitos eléctricos y los digrafos del capítulo 6 los realizaron Miriam Del Valle y Samantha Del Valle. Los planos en tres dimensiones de la figura 1-2 los construyó Elién Rodríguez Del Valle. Ernesto y Miriam hicieron la revisión, en computadora, de las respuestas numéricas de muchos de los ejercicios propuestos y Miriam leyó el texto en su totalidad para localizar erratas. Mi más sincero agradecimiento a todos ellos por la desinteresada ayuda que me brindaron.

Doy gracias a las autoridades del campus Estado de México, del Instituto Tecnológico y de Estudios Superiores de Monterrey, por las facilidades que me dieron para la realización de esta obra; y a Enrique Ortiz, de HP Calculators Latin America, por el soporte otorgado para la realización de la sección 7.1.

El doctor Francisco Delgado Cepeda, profesor del campus Estado de México, leyó por completo el primer capítulo; le agradezco mucho su colaboración y valiosos comentarios.

El doctor Fermín Acosta Magallanes, profesor del campus Estado de México y de la UPIITA del IPN, sacrificó mucho de su tiempo al leer casi en su totalidad el libro. Sus comentarios, observaciones y correcciones fueron de un enorme valor. Obviamente cualquier error técnico en el texto es absolutamente mi responsabilidad. El interés constante que mantuvo Fermín en la realización de esta obra fue un gran estímulo para su culminación y estaré siempre agradecido con él.

Cuando estaba escribiendo este trabajo, se presentaron algunos problemas serios en mi salud, y gracias a los cuidados y apoyo de mis hermanas Rosa María y Gabriela, mi hermano Manuel, mi cuñado José Manuel Lara, mi doctora de cabecera Daniela Lara Del Valle, mis sobrinos Emmanuel Lara, Etzel Rodríguez, Rosa María Lara, Noemí Del Valle, Alejandro Urban, y mis hijas Samantha y Miriam, ahora estoy escribiendo estas últimas líneas. Espero que ellos sepan que pueden contar siempre conmigo como yo conté con ellos.

Escribir un libro, especialmente uno como éste, es una labor en la que hay gran sacrificio no sólo del autor, sino también de los que son más cercanos a él: su familia; en este caso mis hijas Samantha y Miriam. Su paciencia, amor y comprensión fueron el principal incentivo para llegar al final de este proyecto.

Finalmente quiero agradecer a Rubén Dario Santiago Acosta, director del Departamento de Matemáticas y Física del campus Estado de México, por su valiosa cooperación para la realización de este libro. En la vida de todo ser humano existen periodos en que los avatares son más intensos y frecuentes; el lapso para realizar esta obra fue una de esas etapas para mí. Rubén fue en todo momento un apoyo y, aunque la suerte no siempre está de mi lado, soy muy afortunado por tener a un gran amigo como él.

Prólogo

Este libro tiene su germen en las notas del curso semestral de álgebra lineal que he impartido a lo largo de varios años en el Instituto Tecnológico y de Estudios Superiores de Monterrey, campus Estado de México, las cuales son el esqueleto de lo que ahora pretendo mostrar como un cuerpo ya con piel y completo, que se desarrolló gracias a la experiencia adquirida a través de todos esos años.

El objetivo principal es presentar a detalle y profundidad los principales temas del álgebra lineal, mostrando la utilidad de esta materia por medio de una gran variedad de aplicaciones a otros campos y a las propias matemáticas. Integrando la teoría, la práctica, el uso de tecnología y los métodos númericos de esta disciplina.

El libro está diseñado de tal manera que se puede usar para un curso de uno o dos semestres, dependiendo de los programas de estudio de cada institución y de la profundidad con la que se desee tratar cada tema. En el primer caso conviene cubrir las partes I y II, exceptuando los apartados 4.2, 5.3.3 y 5.3.4. Para el segundo caso, se recomiendan todos los temas de las partes I y II, las formas canónicas de Jordan del apéndice D y el material adicional que se incluye en el sitio web del libro. En ambas modalidades se pueden incluir las secciones que se consideren adecuadas de la parte III, especialmente las aplicaciones del capítulo 6.

Como su nombre lo indica, *Álgebra lineal para estudiantes de ingeniería y ciencias* está orientado para ser utilizado tanto en escuelas de ingeniería como en escuelas de ciencias, ya sea a nivel licenciatura o posgrado. Los requisitos académicos para la comprensión del material son las matemáticas elementales que se cubren a *nivel medio superior* (álgebra, geometría analítica y cálculo diferencial e integral).

La mayoría de los estudiantes que toman un curso de álgebra lineal, salvo los que cursan la carrera de matemáticas, se enfrentan por primera vez a una materia en la que se tienen que hacer demostraciones de teoremas y proposiciones matemáticas utilizando el método lógico-deductivo; es la principal dificultad que entraña un curso de esta naturaleza para el lector profano en el campo del rigor matemático. Sin embargo, en álgebra lineal la mayoría de las demostraciones son constructivas; es decir, la prueba de un teorema es en sí un algoritmo para resolver una serie de importantes problemas; lo cual representa una ventaja didáctica para poder iniciarse en el rigor lógico de las matemáticas. Aun tomando en consideración esa ventaja, aprender en qué consiste probar rigurosamente proposiciones matemáticas no es fácil. Para apoyar al estudiante en esta tarea, el apéndice A.2 contiene una breve introducción al método deductivo y a los métodos de demostración en matemáticas —diseñada para que el lector pueda estudiarla por cuenta propia o con un poco de ayuda de su profesor—, a través de casos concretos y con un mínimo de conocimientos previos que seguramente todo estudiante, a este nivel, posee. En estos tiempos, donde la credulidad y las pseudociencias son estimuladas mediáticamente como instrumentos de mercadotecnia para vender productos que “curan” todos los males —incluyendo los políticos y sociales—, el escepticismo, como una cultura de *lo que se afirma se demuestra*, debería ser cultivado por el *Homo sapiens* moderno y el álgebra lineal es una excelente oportunidad para iniciarse, al menos en la parte matemática, en esa cultura.

He dividido el libro en tres partes que, desde mi punto de vista, conforman lo que es el álgebra lineal. Las primeras dos contienen el núcleo teórico de la materia. La parte I —matrices, sistemas lineales, determinantes e inversas de matrices— es la más elemental y es la columna vertebral en la que se apoya el resto del libro; mientras que la II —espacios vectoriales, producto interior, normas, valores y vectores propios— es el corpus de ese núcleo que incluye los temas más relevantes del álgebra lineal. Estos dos segmentos constituyen los primeros cinco capítulos de la obra, y en ellos he intentado exponer el significado matemático del álgebra lineal. En la parte III —que contiene los últimos tres capítulos del texto—, a través de diversas aplicaciones en el capítulo 6, he tratado de hacer patente la utilidad práctica que tiene esta importante materia. Los cálculos numéricos en álgebra lineal pueden llegar a ser muy complejos aritméticamente y tomar demasiado tiempo realizarlos; afortunadamente en esta época contamos con tecnología para apoyarnos en esta tarea. En el capítulo 7, incluí el uso de la tecnología en el álgebra lineal, específicamente con MATLAB y la calculadora HP 50g; y EXCEL, para programación lineal. Sin embargo, una exposición del álgebra lineal que no muestra las dificultades inherentes que se presentan al hacer cálculos numéricos en esta materia y cómo resolverlas matemáticamente, es incompleta. Por esta razón, el capítulo final contiene una introducción relativamente profunda de los principales métodos numéricos que se utilizan en álgebra lineal; con más de 32 programas en MATLAB de esos algoritmos para ser utilizados o modificados, en este u otro lenguaje, por el estudiante a su conveniencia.

Al escribir esta obra intenté tener siempre presentes los obstáculos a los que se enfrentan la mayoría de los estudiantes de álgebra lineal, el principal es el alto nivel de abstracción de la materia. Para soslayar esta dificultad, el libro contiene más de 200 figuras con el propósito de crear imágenes que puedan ayudar al lector a visualizar física y geométricamente entes abstractos y convertirlos en conceptos más concretos. Además, a lo largo de sus 8 capítulos y 5 apéndices, incluí más de 450 ejemplos para apoyarlo a comprender la materia. Sin embargo, pensé que esto no era suficiente, pues el estudiante necesita ver cómo se resuelven ejercicios en álgebra lineal, sobre todo aquellos que tienen contenidos altos de abstracción; por esta razón incorporé, en la última sección de cada uno de los primeros cinco capítulos —que conforman el núcleo principal del libro— un grupo de ejercicios resueltos con detalle; en total forman un conjunto de más de 230 ejercicios de cálculos directos, demostraciones, etc., que junto con los ejemplos del texto suman un total de más de 680 problemas completamente resueltos que el lector puede consultar según lo necesite. Naturalmente, no basta con “ver”, se necesita “hacer” y, para ello, el libro contiene —al final de cada capítulo— una sección de ejercicios propuestos al estudiante —con respuestas a los ejercicios seleccionados en el apéndice E— para que practique a discreción o de acuerdo con las instrucciones de su profesor; en total el libro cuenta con más de 2300 ejercicios propuestos.

Con el propósito de no interrumpir la exposición de la teoría en el texto y para facilitar su consulta, coloqué aparte, en el capítulo 6, las aplicaciones. Al principio de cada una de ellas se hacen explícitos los requisitos —del material del texto y de otras disciplinas— que se necesitan para su estudio. El nivel de las aplicaciones aumenta gradualmente desde el muy elemental hasta un nivel que demanda mucho más esfuerzo para su comprensión; sin embargo, confío que la utilidad final que el estudiante encuentre en ellas bien valdrá la pena el tiempo invertido para su estudio. De hecho, este capítulo se puede abordar inmediatamente después de que se cumplan los requisitos que señala la aplicación correspondiente; por ejemplo, las aplicaciones de las secciones 6.1, 6.2, 6.4, 6.5 y 6.6, se pueden tratar en seguida que se ha cubierto el material de matrices y sistemas lineales (o en forma simultánea). Sin embargo, en el texto hay algunas aplicaciones que en realidad están concatenadas a la teoría —por ejemplo, el tema de aproximación óptima en espacios normados, o la interesante teoría de detección y corrección de errores en códigos binarios que está al final del apéndice B—, esas no las incluí en el capítulo 6 y se encuentran dispersas a lo largo del libro; en la página 1109 hay una lista de ellas con referencias al lugar donde se

localizan en el texto. Una función semejante cumple el listado de la página 1110, que es una descripción de los principales programas en MATLAB que contiene el libro y señala su ubicación.

Además, esta obra cuenta con una página donde el estudiante tendrá acceso a diversos recursos: www.mhhe.com/uni/delvalleag1e.

Espero que *Álgebra lineal para estudiantes de ingeniería y ciencias* cumpla con los propósitos para los que fue creado, sirva de apoyo a la labor docente de los profesores que trabajan educando en esta materia y que vosotros, estudiantes, encuentren en él no sólo dónde aprender álgebra lineal, sino que también disfruten de ese proceso como yo lo hice al escribir cada una de las líneas de este libro (también sufrí, ojalá ustedes no).

México D.F., primavera de 2011

JUAN CARLOS DEL VALLE SOTELO

I

Matrices, sistemas y determinantes

1 Matrices y sistemas lineales

En este capítulo se introducen los conceptos básicos que se requieren para estudiar álgebra lineal. Comenzamos en la primera sección con el tema fundamental de matrices. Las matrices se crearon para operar ciertos arreglos numéricos que aparecen tanto en aplicaciones como en las propias matemáticas. Continuamos en la segunda sección con el estudio general de sistemas de ecuaciones lineales. Los sistemas de ecuaciones lineales tienen una gran variedad de aplicaciones en ciencias e ingeniería y seguramente el lector ya tuvo algún contacto con ellos en forma elemental en secundaria y bachillerato; aquí nos abocamos a un estudio general y profundo de este importante tema. La tercera sección contiene un compendio de ejercicios resueltos de las dos secciones precedentes para que el lector consulte el mayor número de ejemplos resueltos y un conjunto de ejercicios propuestos para que los resuelva el estudiante.

1.1 Matrices

1.1.1 Definiciones y ejemplos

Definición 1.1 *Una matriz A es un arreglo de m-renglones o filas y n-columnas de $m \times n$ números reales:*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Se dice entonces que A es una matriz de tamaño $m \times n$ y simbólicamente se escribe

$$A = [a_{ij}]\,,$$

$i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$. Esto es, a_{ij} representa el número que se encuentra en la fila i y en la columna j. A los elementos a_{ij} se les llaman las componentes (entradas) de la matriz A.

❍ **Nota 1.1**

1. Los paréntesis rectangulares se pueden suplir por paréntesis circulares en notaciones matriciales. En este libro emplearemos paréntesis rectangulares.

2. En el caso particular de que una matriz tenga tamaño 1×1 escribiremos simplemente a en lugar de $[a]$; es decir, identificaremos toda matriz $[a]$ con el número real a.

▶ **Ejemplo 1.1** Si

$$A=\begin{bmatrix}-2 & 3 & 5\\ -4 & 2 & 1\end{bmatrix},$$

A es una matriz 2×3 y, para este caso, $a_{11}=-2$, $a_{12}=3$, $a_{13}=5$, $a_{21}=-4$, $a_{22}=2$, $a_{23}=1$.◀

❍ **Nota 1.2** Al conjunto de matrices de tamaño $m\times n$ lo denotaremos, en este libro, por $\mathfrak{M}_{m\times n}$.

Definición 1.2 *Dos matrices* $A=[a_{ij}]$, $B=[b_{ij}]$ **son iguales** *(*$A=B$*) si y sólo si:*

- *A y B tienen el mismo tamaño y*
- $a_{ij}=b_{ij}\ \ \forall i,j.$

▶ **Ejemplo 1.2** De acuerdo con la definición precedente $\begin{bmatrix}1 & 3 & 9\\ 5 & 7 & 2\end{bmatrix}\neq\begin{bmatrix}1 & 3 & 9\\ 5 & 6 & 2\end{bmatrix}$. ◀

▶ **Ejemplo 1.3** Determinar el valor de a para que las matrices $A=\begin{bmatrix}a & 1\\ -1 & 2a\end{bmatrix}$ y $B=\begin{bmatrix}2 & 1\\ -1 & 4\end{bmatrix}$ sean iguales.◀

Solución Dado que ambas matrices tienen el mismo tamaño ellas serán iguales si y sólo si coinciden componente a componente; para lo cual es suficiente que $a=2$ y $2a=4$; esto es, para $a=2$. ✔

▶ **Ejemplo 1.4** Resolver el ejemplo anterior si $A=\begin{bmatrix}a & 0\\ 3 & 3a\end{bmatrix}$ y $B=\begin{bmatrix}1 & 0\\ 3 & 4\end{bmatrix}$.◀

Solución Para que las matrices sean iguales se requiere, en este caso, que $a=1$ y $3a=4$, luego se debe tener simultáneamente $a=1$ y $a=4/3$; lo cual es imposible. Por tanto $A\neq B$ para cualquier valor de a. ✔

1.1.2 Operaciones con matrices

1. **Multiplicación de un escalar**[1] **con una matriz**. Si $\lambda\in\mathbb{R}$ y $A=[a_{ij}]\in\mathfrak{M}_{m\times n}$ se define $\lambda A=[\lambda a_{ij}]$. Es decir, el resultado de multiplicar una matriz con un escalar es la matriz que tiene como componentes cada una de las entradas de la matriz original multiplicada por dicho escalar.
2. **Suma de matrices**. Si $A,B\in\mathfrak{M}_{m\times n}$, $A=[a_{ij}]$, $B=[b_{ij}]$; se define la suma de A con B como $A+B=[c_{ij}]$, con $c_{ij}=a_{ij}+b_{ij}\ \forall i,j$. Así, la suma de dos matrices sólo se puede realizar cuando éstas tienen el mismo tamaño y el resultado es también una matriz $m\times n$.

[1] Diremos que todo número real es un escalar.

3. **Multiplicación de una matriz fila por matriz columna.**[2]

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ \cdot \\ \cdot \\ \cdot \\ b_{n1} \end{bmatrix} = a_{11}b_{11} + a_{12}b_{21} + \cdots a_{1n}b_{n1}.$$

De acuerdo con esta definición, el producto de una matriz fila con una matriz columna sólo se puede llevar a cabo cuando la primera tiene tamaño $1 \times n$ y la segunda $n \times 1$ (las dos tienen el mismo número de componentes) y el resultado de la operación será una matriz 1×1 (un número real).

4. **Producto de una matriz $m \times n$ con una matriz $n \times p$.** Si $A = [a_{ij}] \in \mathfrak{M}_{m \times n}$ y $B = [b_{ij}] \in \mathfrak{M}_{n \times p}$, el producto de A con B se define como $AB = [c_{ij}]$ donde

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj},$$

para $i = 1, 2, \ldots, m$ y $j = 1, 2, \ldots, p$. Es decir, la componente c_{ij} del producto AB es el resultado de multiplicar la i-ésima fila de A con la j-ésima columna de B. Además, para poder efectuar el producto, la primera matriz debe tener el mismo número de columnas que de filas la segunda y la matriz AB tiene entonces tamaño $m \times p$. En forma equivalente, si F_i, $i = 1, \ldots, m$, son las filas de A y C_j, $j = 1, \ldots, p$, son las columnas de B, entonces

$$AB = \begin{bmatrix} F_1C_1 & F_1C_2 & \cdots & F_1C_p \\ F_2C_1 & F_2C_2 & \cdots & F_1C_p \\ \vdots & \vdots & \ddots & \vdots \\ F_mC_1 & F_2C_2 & \cdots & F_mC_p \end{bmatrix} \qquad (1.1)$$

► **Ejemplo 1.5**

- $\sqrt{2}\begin{bmatrix} -1 & 0 & -1 & 2 \\ 2 & -4 & 1 & 3 \\ \sqrt{2} & -4 & 0 & 5 \end{bmatrix} = \begin{bmatrix} -\sqrt{2} & 0 & -\sqrt{2} & 2\sqrt{2} \\ 2\sqrt{2} & -4\sqrt{2} & \sqrt{2} & 3\sqrt{2} \\ 2 & -4\sqrt{2} & 0 & 5\sqrt{2} \end{bmatrix}$

- Si $A = \begin{bmatrix} -2 & -4 & -1 \\ 5 & -2 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} -4 & -5 & 2 \\ -1 & 0 & -1 \end{bmatrix}$, entonces $A + B = \begin{bmatrix} -6 & -9 & 1 \\ 4 & -2 & -1 \end{bmatrix}$.

- $$\begin{bmatrix} -1 & 0 & -2 & 4 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 0 \\ 0 \\ -4 \end{bmatrix} = (-1)(2) + (0)(-1) + (-2)(0) + (4)(0) + (5)(-4) = -22.$$

Note que en este caso la matriz fila tiene tamaño 1×5 y la columna 5×1 (las dos tienen el mismo número de componentes).◄

[2] Una matriz fila es una matriz que tiene solamente un renglón y una matriz columna es una matriz que tiene una sola columna (cfr. inciso 3 de la pág. 8).

▶ **Ejemplo 1.6** Si

$$A = \begin{bmatrix} -1 & -2 & 4 \\ 0 & 2 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & -2 & 4 & 5 \\ 0 & -1 & 0 & 2 \\ -1 & 0 & 0 & 1 \end{bmatrix},$$

$A \in \mathfrak{M}_{2\times 3}$, $B \in \mathfrak{M}_{3\times 4}$; el producto AB está definido (el número de columnas de A es igual al número de filas de B, en este caso 3) y el producto AB será una matriz 2×4, dos filas y cuatro columnas (tantas filas como A y tantas columnas como B). Para obtener las componentes c_{ij} de las filas de la matriz producto AB procedemos de la manera siguiente.

La primera fila de AB: Los elementos de la primera fila de AB se obtienen multiplicando, sucesivamente, la primera fila de A con la primera, segunda, tercera y cuarta columas de B:

$$c_{11} = \begin{bmatrix} -1 & -2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = -5,$$

$$c_{12} = \begin{bmatrix} -1 & -2 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix} = 4,$$

$$c_{13} = \begin{bmatrix} -1 & -2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} = -4,$$

$$c_{14} = \begin{bmatrix} -1 & -2 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix} = -5.$$

La segunda fila de AB: Los elementos de la segunda fila de AB se obtienen multiplicando, sucesivamente, la segunda fila de A con la primera, segunda, tercera y cuarta columas de B:

$$c_{21} = \begin{bmatrix} 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = -1,$$

$$c_{22} = \begin{bmatrix} 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix} = -2,$$

$$c_{23} = \begin{bmatrix} 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} = 0,$$

$$c_{24} = \begin{bmatrix} 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix} = 5.$$

Luego,

$$AB = \begin{bmatrix} -5 & 4 & -4 & -5 \\ -1 & -2 & 0 & 5 \end{bmatrix}. ◀$$

En realidad, la notación matricial está diseñada para ejecutar mecánica y mentalmente los cálculos cuando el tamaño de las matrices no es muy grande; por eso el lector debe procurar, en la medida de lo posible, aprovechar esta ventaja para efectuar las operaciones de esta manera. De hecho, a partir de aquí, el lector ya no encontrará un producto de matrices realizado con el detalle con el que se hizo en el ejemplo precedente; pues utilizaremos sistemáticamente (1.1) para producto de matrices y haremos los cálculos sin hacer explícitas las operaciones.

► **Ejemplo 1.7**

$$\begin{aligned}\begin{bmatrix}-1 & 0 & 1\\ 2 & 1 & 1\\ 3 & -2 & 0\end{bmatrix}\begin{bmatrix}0 & -1 & 1\\ 1 & 1 & -1\\ 0 & 1 & 2\end{bmatrix} &= \begin{bmatrix}F_1C_1 & F_1C_2 & F_1C_3\\ F_2C_1 & F_2C_2 & F_2C_3\\ F_3C_1 & F_3C_2 & F_3C_3\end{bmatrix}\\ &= \begin{bmatrix}0 & 2 & 1\\ 1 & 0 & 3\\ -2 & -5 & 5\end{bmatrix}. \blacktriangleleft\end{aligned}$$

1.1.3 Matrices especiales

1. Matriz cero. La matriz cero de tamaño $m\times n$ se define como aquella que tiene las $m\times n$ componentes nulas; esto es,

$$\mathscr{O} = [a_{ij}]$$

donde $a_{ij} = 0\ \forall i, j$. Así, por ejemplo,

$$\mathscr{O} = \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$$

es la matriz cero 2×3.

2. Matriz identidad $n\times n$:

$$I_n = \begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix};$$

es decir, $I_n = [a_{ij}]$, donde

$$a_{ij} = \begin{cases}1, & \text{si } i = j;\\ 0, & \text{si } i \neq j.\end{cases}$$

Así, por ejemplo,

$$I_3 = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

es la matriz identidad 3×3.

3. Como mencionamos en el inciso 3 de la subsección 1.1.2, a las matrices que tienen sólo una fila o sólo una columna les llamaremos, respectivamente, **matrices fila** y **matrices columna**. Además, en este libro utilizaremos una notación especial en el caso de las matrices columna (cuando tengan más de un elemento) análoga a la notación vectorial

$$\vec{b} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}.$$

La razón de esta notación se verá más adelante cuando se estudie el espacio vectorial $\mathbb{R}^n$ en el capítulo 3.

A las matrices de tamaño $n \times n$ les llamaremos **matrices cuadradas** de orden n y al conjunto formado por éstas lo denotaremos por $\mathfrak{M}_n$. Si $A = [a_{ij}]$ es una matriz cuadrada de orden n se dice que los elementos a_{11}, a_{22}, a_{33},..., a_{nn} forman o están en la **diagonal** de la matriz A. Y si $A = [a_{ij}] \in \mathfrak{M}_{m\times n}$, diremos que los elementos a_{ij} con $i = j$ forman la **diagonal principal** de la matriz A.

► **Ejemplo 1.8** Si

$$M = \begin{bmatrix} -1 & 5 & 0 & 2 \\ 7 & 3 & -1 & 1 \\ 3 & 0 & 4 & 2 \\ 1 & -5 & 9 & 7 \end{bmatrix}$$

entonces $m_{11} = -1, m_{22} = 3, m_{33} = 4, m_{44} = 7$ son los elementos de la diagonal de la matriz cuadrada M.◄

Definición 1.3 *Una matriz cuadrada A de orden n es* **triangular superior** *si las componentes que están por debajo de la diagonal son todas nulas. La matriz es* **triangular inferior** *si las componentes que están por arriba de la diagonal son todas iguales a cero.*

► **Ejemplo 1.9** Si

$$A = \begin{bmatrix} -1 & 5 & 0 & 2 \\ 0 & 3 & -1 & 1 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 7 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} -1 & 0 & 0 & 0 \\ -5 & 3 & 0 & 0 \\ 2 & 0 & 4 & 0 \\ 6 & 0 & 4 & 0 \end{bmatrix},$$

entonces A es una matriz triangular superior y B es una matriz triangular inferior.◄

Definición 1.4 *Sea A una matriz cuadrada de orden n. Se dice que A es una matriz* **diagonal** *si todas las componentes fuera de su diagonal son nulas. Si* $a_{ii} = \lambda_i$, $i = 1, 2, \ldots, n$, *son las componentes de la diagonal de esta matriz se escribe*

$$A = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

para representar a la matriz diagonal A.

▶ **Ejemplo 1.10** La matriz cuadrada $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 8 \end{bmatrix}$ es diagonal. Esto es,

$$A = \operatorname{diag}(4,3,8).$$ ◀

Definición 1.5 *Si $A = [a_{ij}] \in \mathfrak{M}_{m\times n}$ se define la* **matriz transpuesta** *de A como $A^t = [b_{ij}]$, donde $b_{ij} = a_{ji}$ para $i = 1,2,\ldots,n$ y $j = 1,2,\ldots,m$.*

De la definición 1.5 se desprende que A^t tiene tamaño $n \times m$ y que en la matriz transpuesta la primera columna es la primera fila de A, la segunda columna es la segunda fila de A, etcétera.

Definición 1.6 *Una matriz A es* **simétrica** *cuando $A^t = A$.*

La definición 1.6 entraña que una matriz simétrica es necesariamente cuadrada; pues si $A \in \mathfrak{M}_{m\times n}$ y A es simétrica, entonces $A = A^t \in \mathfrak{M}_{n\times m}$, de donde $m = n$; ya que dos matrices que son iguales deben tener el mismo tamaño.

▶ **Ejemplo 1.11** Si

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix},$$

$$A^t = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}.$$ ◀

▶ **Ejemplo 1.12** La matriz

$$A = \begin{bmatrix} -1 & 2 \\ 2 & 3 \end{bmatrix}$$

es simétrica pues claramente $A = A^t$. ◀

1.1.4 Propiedades de las operaciones

A continuación enunciamos las principales propiedades de las operaciones con matrices, las cuales son, en general, fáciles de probar y su comprobación se deja como ejercicio al lector.

1. Si $A, B, C \in \mathfrak{M}_{m\times n}$ y $\lambda, \beta \in \mathbb{R}$:

   (a) $A + B \in \mathfrak{M}_{m\times n}$.

   (b) $A + (B + C) = (A + B) + C$.

(*c*) $A+B=B+A$.

(*d*) $A+\mathscr{O}=A$, donde $\mathscr{O}$ es la matriz cero $m\times n$.

(*e*) Existe una matriz $-A\in\mathfrak{M}_{m\times n}$ tal que $A+(-A)=\mathscr{O}$. De hecho, si $A=[a_{ij}]$, $-A=[-a_{ij}]$.

(*f*) $\lambda A\in\mathfrak{M}_{m\times n}$.

(*g*) $\lambda(\beta A)=(\lambda\beta)A$.

(*h*) $(\lambda+\beta)A=\lambda A+\beta A$.

(*i*) $\lambda(A+B)=\lambda A+\lambda B$.

(*j*)[3] $1A=A$.

2. (*a*) Si A, B, C son matrices tales que los productos $A(BC)$ y $(AB)C$ están definidos, entonces

$$A(BC)=(AB)C.$$

(*b*) Si AB está definido se tiene: $\lambda(AB)=(\lambda A)B=A(\lambda B)$.

(*c*) Si $A\in\mathfrak{M}_{m\times n}$, $AI_n=I_mA=A$.

(*d*) En general $AB\neq BA$.

(*e*) Si $A\in\mathfrak{M}_{m\times n}$ y $B,C\in\mathfrak{M}_{n\times p}$, entonces

$$A(B+C)=AB+AC.$$

3. (*a*) Si A y B son matrices del mismo tamaño $(A+B)^t=A^t+B^t$.

(*b*) Si A, B son matrices tales que el producto AB está definido, entonces $(AB)^t=B^tA^t$.

(*c*) $(A^t)^t=A\ \forall A\in\mathfrak{M}_{m\times n}$.

Es conveniente que el lector tenga siempre presente la propiedad 2(*d*); es decir, la **no conmutatividad** del producto de matrices. Pues es claro que en principio el hecho de que el producto AB esté definido, no garantiza que ni siquiera el producto BA esté definido; por ejemplo, si A es una matriz 2×3 y B es una matriz 3×4, el producto AB está definido y el producto BA no. Más aún, aunque los productos AB y BA estén definidos éstos, en general, serán distintos como ilustramos en el siguiente ejemplo.

▶ Ejemplo 1.13

$$\begin{bmatrix}1&1\\3&2\end{bmatrix}\begin{bmatrix}1&0\\2&4\end{bmatrix}=\begin{bmatrix}3&4\\7&8\end{bmatrix},$$

$$\begin{bmatrix}1&0\\2&4\end{bmatrix}\begin{bmatrix}1&1\\3&2\end{bmatrix}=\begin{bmatrix}1&1\\14&10\end{bmatrix};$$

esto es,

$$\begin{bmatrix}1&1\\3&2\end{bmatrix}\begin{bmatrix}1&0\\2&4\end{bmatrix}\neq\begin{bmatrix}1&0\\2&4\end{bmatrix}\begin{bmatrix}1&1\\3&2\end{bmatrix}$$ ◀

Finalizamos este apartado con las demostraciones, en los siguientes dos ejemplos, de un par de propiedades simples del producto de matrices que serán utilizadas más adelante.

[3]Más adelante, en el tema de espacios vectoriales, se verá la importancia de esta aparentemente inocua propiedad.

▶ **Ejemplo 1.14** Sean $A = [a_{ij}] \in \mathfrak{M}_{m\times n}$ y $C = [b_{ij}] \in \mathfrak{M}_{n\times p}$. Si $\vec{c}_k = \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix}$ es la columna k de C y $\vec{d}_k$ es la columna k de AC, $k = 1, 2, \ldots, p$, demostrar que

$$\vec{d}_k = A\vec{c}_k \quad \forall k.$$

Esto es,

$$AC = \begin{bmatrix} A\vec{c}_1 & A\vec{c}_2 & \cdots & A\vec{c}_p \end{bmatrix} \blacktriangleleft \tag{1.2}$$

DEMOSTRACIÓN ■ Sean α_{ij} las componentes del producto AC, entonces, para cada $k = 1, 2, \ldots, p$,

$$\vec{d}_k = \begin{bmatrix} \alpha_{1k} \\ \alpha_{2k} \\ \vdots \\ \alpha_{mk} \end{bmatrix};$$

pero

$$\begin{aligned} \alpha_{ik} &= \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix} \\ &= \sum_{j=1}^{n} a_{ij} b_{jk}; \end{aligned}$$

por tanto,

$$\vec{d}_k = \begin{bmatrix} \sum_{j=1}^{n} a_{1j} b_{jk} \\ \sum_{j=1}^{n} a_{2j} b_{jk} \\ \vdots \\ \sum_{j=1}^{n} a_{mj} b_{jk} \end{bmatrix}. \tag{1.3}$$

Por otra parte,

$$\begin{aligned} A\vec{c}_k &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{j=1}^{n} a_{1j} b_{jk} \\ \sum_{j=1}^{n} a_{2j} b_{jk} \\ \vdots \\ \sum_{j=1}^{n} a_{mj} b_{jk} \end{bmatrix}. \end{aligned} \tag{1.4}$$

De (1.3) y (1.4) se tiene $A\vec{c}_k = \vec{d}_k \; \forall k$. ■

► **Ejemplo 1.15** Supongamos ahora que $A = [a_{ij}] \in \mathfrak{M}_{m\times n}$ y $\vec{c} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, entonces,

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{1.5}$$

En efecto:

$$\begin{aligned} x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} &= \begin{bmatrix} x_1a_{11} \\ x_1a_{21} \\ \vdots \\ x_1a_{m1} \end{bmatrix} + \begin{bmatrix} x_2a_{12} \\ x_2a_{22} \\ \vdots \\ x_2a_{m2} \end{bmatrix} + \cdots + \begin{bmatrix} x_na_{1n} \\ x_na_{2n} \\ \vdots \\ x_na_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \\ &= A\vec{c}. \end{aligned}$$ ◄

1.1.5 Matrices con números complejos

En este apartado se introduce, por primera vez en este libro, el uso de números complejos en álgebra lineal; específicamente en el tema de matrices con componentes complejas. El ápendice B contiene un breve estudio de este importante campo numérico y de sus principales propiedades, y el lector que no esté habituado a trabajar con números complejos, o necesite repasar este tema, debería consultar la sección B.1 de este apéndice cuanto antes. A lo largo de este texto se incluyen apartados que contienen el uso de números complejos en temas que ya se han tratado con números reales. En general, la transición en cada caso será muy sencilla, pues una vez que se dominan los temas de álgebra lineal con números reales los cambios para tratar éstos con números complejos son mínimos y, en realidad, las dificultades tienen que ver más con la familiaridad que tenga el lector con el uso de números complejos que con aspectos áridos de generalización. De hecho, el uso de este campo numérico en álgebra lineal se va haciendo cada vez más necesario en la medida que se avanza en la materia, tanto en la teoría como en las aplicaciones. Se han incluido este tipo de subsecciones para ir acostumbrando al lector al uso de los números complejos en álgebra lineal. Obviamente, el lector que no desee en este momento abordar estos temas puede omitirlos y regresar a ellos cuando lo juzgue pertinente.

Recordemos (cfr. apéndice B) que los números complejos tienen la forma

$$a + bi$$

donde a, b son números reales e i es la unidad *imaginaria*. Al conjunto de estos números se les representa por $\mathbb{C}$ y este campo incluye de manera natural a los números reales mediante la identificación del número real a con el número complejo $a+0i$. Estos números se operan algebraicamente de manera análoga a los números reales, utilizando todas las propiedades de éstos y conviniendo en que la unidad *imaginaria* en este sistema satisface[4]

$$i^2 = -1.$$

De esta manera, operar algebraicamente matrices con componentes complejas es un proceso completamente análogo al que se utiliza cuando éstas tienen entradas que son números reales. Es decir, se suman, restan, multiplican, etc., en la misma forma que las matrices reales, pero operando sus componentes con las reglas algebraicas de los números complejos. Al conjunto de matrices de tamaño $m \times n$ con componentes complejas lo denotaremos por $\mathfrak{M}_{m\times n}(\mathbb{C})$. Todas las propiedades acerca de matrices con componentes reales que vimos en esta sección siguen siendo válidas para las matrices con entradas complejas.

► **Ejemplo 1.16** Sean $A, B \in \mathfrak{M}_{2\times 3}(\mathbb{C})$ las matrices definidas por

$$A = \begin{bmatrix} 1-2i & -4i & 2 \\ 3-5i & 4+6i & -9i \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 3-7i & 5-4i & 2-9i \\ 5i & 7-6i & 1+i \end{bmatrix}.$$

Entonces

1. $$\begin{aligned} A+B &= \begin{bmatrix} 1-2i & -4i & 2 \\ 3-5i & 4+6i & -9i \end{bmatrix} + \begin{bmatrix} 3-7i & 5-4i & 2-9i \\ 5i & 7-6i & 1+i \end{bmatrix} \\ &= \begin{bmatrix} 4-9i & 5-8i & 4-9i \\ 3 & 11 & 1-8i \end{bmatrix}. \end{aligned}$$

2. $$\begin{aligned} 5A &= 5\begin{bmatrix} 1-2i & -4i & 2 \\ 3-5i & 4+6i & -9i \end{bmatrix} \\ &= \begin{bmatrix} 5-10i & -20i & 10 \\ 15-25i & 20+30i & -45i \end{bmatrix}. \end{aligned}$$

3. $$\begin{aligned} (3+2i)B &= (3+2i)\begin{bmatrix} 1-2i & -4i & 2 \\ 3-5i & 4+6i & -9i \end{bmatrix} \\ &= \begin{bmatrix} 7-4i & 8-12i & 6+4i \\ 19-9i & 26i & 18-27i \end{bmatrix}. \end{aligned}$$

   Aquí hemos realizado las operaciones

$$\begin{aligned} (3+2i)(1-2i) &= 3-6i+2i-4i^2 \\ &= 3-4i-4(-1) \\ &= 3-4i+4 \\ &= 7-4i, \end{aligned}$$

[4] En la sección B.1 del apéndice B se hace un estudio más detallado y formal de los números complejos.

para obtener la componente c_{11} de $(3+2i)B$;

$$\begin{aligned}(3+2i)(-4i) &= -12i-8i^2\\ &= -12i-8(-1)\\ &= 8-12i,\end{aligned}$$

para obtener la componente c_{12} de $(3+2i)B$; etcétera.◄

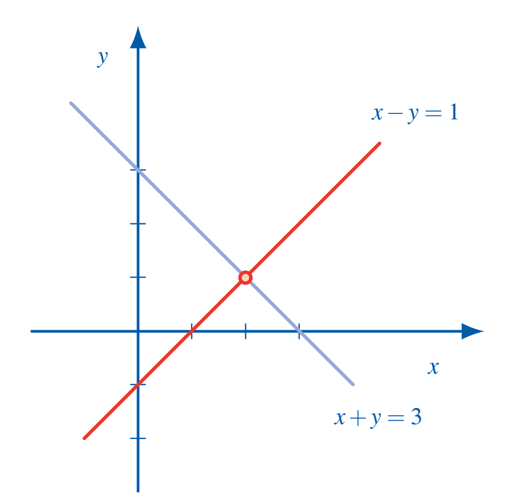
► **Ejemplo 1.17** Sean

$$A=\begin{bmatrix}1+i & 2\\ -i & 2-3i\end{bmatrix} \quad \text{y} \quad B=\begin{bmatrix}-i & 3 & 2+5i\\ 2i & 1-i & 0\end{bmatrix},$$

entonces

$$\begin{aligned}AB &= \begin{bmatrix}1+i & 2\\ -i & 2-3i\end{bmatrix}\begin{bmatrix}-i & 3 & 2+5i\\ 2i & 1-i & 0\end{bmatrix}\\ &= \begin{bmatrix}(1+i)(-i)+2(2i) & (1+i)(3)+2(1-i) & (1+i)(2+5i)+2(0)\\ (-i)(-i)+(2-3i)(2i) & (-i)(3)+(2-3i)(1-i) & (-i)(2+5i)+(2-3i)(0)\end{bmatrix}\\ &= \begin{bmatrix}1+3i & 5+i & -3+7i\\ 5+4i & -1-8i & 5-2i\end{bmatrix}.\end{aligned}$$ ◄

1.2 Sistemas lineales



Seguramente el lector está familiarizado, por cursos más elementales, con sistemas simultáneos de dos o tres ecuaciones lineales con dos o tres incógnitas. Se les llama sistemas lineales porque, para el caso de dos incógnitas, digamos x, y, las ecuaciones tienen la forma $ax+by=c$, cuyos lugares geométricos corresponden a líneas rectas en el plano. Cuando se resuelve un sistema lineal de dos ecuaciones con dos incógnitas, se busca el punto de intersección de dos líneas rectas (si es que éstas no son paralelas). Aquí estudiaremos sistemas lineales generales de m ecuaciones con n incógnitas siendo m y n cualquier par de números enteros no negativos. Los sistemas lineales tienen una gran variedad de aplicaciones en ingeniería y ciencias; veremos algunas de estas aplicaciones en el capítulo seis.

1.2.1 Definiciones, soluciones y forma matricial de sistemas lineales

Definición 1.7 *Un sistema de m-ecuaciones con n-incógnitas que tiene la forma*

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ . & . & . & . & \cdots & . & . & . & . \\ . & . & . & . & \cdots & . & . & . & . \\ . & . & . & . & \cdots & . & . & . & . \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \tag{1.6}$$

donde los $a_{ij}, b_i \in \mathbb{R}$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, están dados, es **lineal**. *Una solución de este sistema de ecuaciones es una n-ada ordenada $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ de números reales, tales que al hacer las sustituciones*

$$\begin{aligned} x_1 &= \alpha_1 \\ x_2 &= \alpha_2 \\ &\vdots \\ x_n &= \alpha_n \end{aligned}$$

en cada una de las m-ecuaciones las convierte en identidades.

► **Ejemplo 1.18** El sistema de dos ecuaciones con tres incógnitas

$$2x_1 - 3x_2 - x_3 = -4 \tag{1.7}$$

$$x_1 + x_2 + x_3 = -3 \tag{1.8}$$

es lineal y $(-1, 2, -4)$ es una solución del mismo. En efecto, al sustituir $x_1 = -1$, $x_2 = 2$ y $x_3 = -4$ en la primera ecuación (1.7) se tiene

$$2(-1) - 3(2) - (-4) = -4$$

y al hacer las mismas sustituciones en la segunda ecuación (1.8),

$$(-1) + (2) + (-4) = -3. \blacktriangleleft$$

► **Ejemplo 1.19** El sistema de dos ecuaciones con dos incógnitas

$$\begin{aligned} x_1^2 - 3x_2 &= 1 \\ x_1^{1/2} + x_2 &= \pi \end{aligned}$$

no es lineal (¿por qué?).◄

Si se tiene el sistema lineal (1.6) a

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

se le llama la **matriz de coeficientes** del sistema. En tal caso, si ponemos

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \text{ y } \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix},$$

entonces el sistema lineal se puede escribir en forma matricial como

$$A\vec{x} = \vec{b},$$

pues al hacer el producto se obtiene

$$\begin{bmatrix} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix}$$

que equivale, por definición de igualdad de matrices, al sistema (1.6).

► **Ejemplo 1.20** Para el sistema 3×3

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & 2x_3 & = & 9 \\ 2x_1 & + & 4x_2 & - & 3x_3 & = & 1 \\ 3x_1 & + & 6x_2 & - & 5x_3 & = & 0 \end{array}$$

la matriz de coeficientes es

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & -3 \\ 3 & 6 & -5 \end{bmatrix}$$

y la ecuación matricial correspondiente es

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & -3 \\ 3 & 6 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 1 \\ 0 \end{bmatrix}. ◄$$

Definición 1.8 *El sistema $m \times n$ $A\vec{x} = \vec{b}$ es:*

- **Consistente**: *si tiene al menos una solución.*
- **Inconsistente**: *si no tiene soluciones.*

En la figura 1-1 se ilustran los lugares geométricos de cuatro sistemas lineales en el plano: con solución única *(a)*, inconsistentes *(b)* y *(c)* y con una infinidad de soluciones *(d)*.

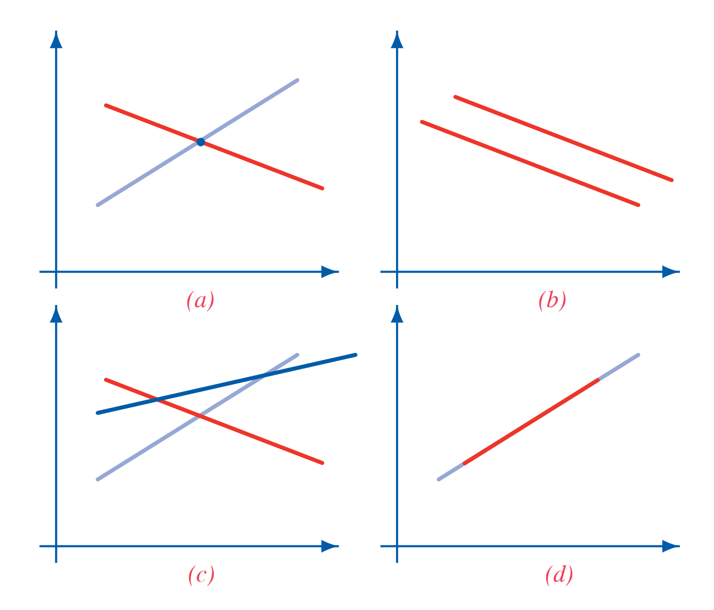


Figura 1-1 • *(a)* dos líneas que se intersecan en un solo punto, *(b)* dos líneas paralelas que no se intersecan, *(c)* tres líneas que no se intersecan simultáneamente y *(d)* dos líneas que coinciden.

De manera análoga, una ecuación lineal con tres incógnitas, $ax + by + cz = d$, corresponde al lugar geométrico de puntos que están en un plano en el espacio tridimensional. También en este caso, cuando se resuelven sistemas lineales con tres incógnitas, se buscan intersecciones de los correspondientes planos. Nuevamente los planos pueden no intersecarse, intersecarse en una infinidad de puntos o intersecarse en un único punto. La figura 1-2 ilustra estas posibilidades.

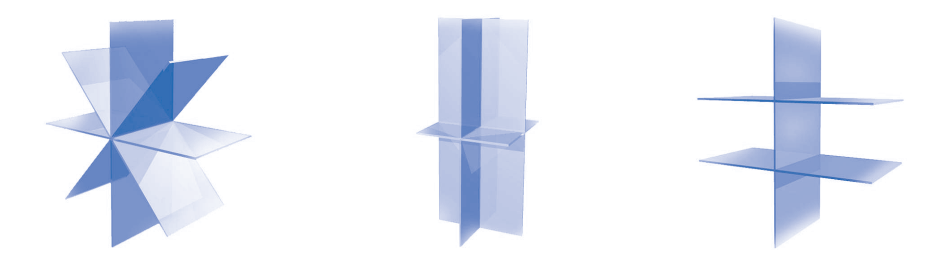


Figura 1-2 • Planos que se intersecan, respectivamente, en una línea recta, en un único punto y que no tienen intersección simultánea.

Definición 1.9 *Dos sistemas lineales del mismo tamaño, $A\vec{x} = \vec{b}$, $H\vec{x} = \vec{c}$, son* **equivalentes** *si tienen el mismo conjunto de soluciones.*

En el siguiente ejemplo resolveremos un sistema lineal de manera análoga a como el lector, seguramente, ya lo ha hecho en cursos de bachillerato; sin embargo, lo haremos con un método que introducirá el importante algoritmo de Gauss; el cual consiste, esencialmente, en ir haciendo “pivotes” para eliminar variables (incógnitas) y obtener un sistema equivalente en forma “escalonada” y finalmente resolverlo por sustitución regresiva.

► **Ejemplo 1.21** Resolvamos el sistema lineal

$$x_1 + x_2 + 2x_3 = 9 \tag{1.9}$$
$$2x_1 + 4x_2 - 3x_3 = 1 \tag{1.10}$$
$$3x_1 + 6x_2 - 5x_3 = 0 \tag{1.11}$$

Para ello, con la ecuación (1.9), eliminemos la variable x_1 de las ecuaciones (1.10) y (1.11) multiplicando[5] (1.9) por -2 y sumando con (1.10); luego multiplicando (1.9) por -3 y sumando con (1.11); obteniendo el sistema equivalente:

$$x_1 + x_2 + 2x_3 = 9$$
$$2x_2 - 7x_3 = -17 \tag{1.12}$$
$$3x_2 - 11x_3 = -27 \tag{1.13}$$

De manera análoga, multiplicando (1.12) por -3, (1.13) por 2 y sumando los resultados, habremos hecho un "pivote" con la variable x_2 de la ecuación (1.12) para eliminar la variable x_2 de la ecuación (1.13), produciendo el sistema equivalente "escalonado"

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 9 \\ & & 2x_2 & - & 7x_3 & = & -17 \\ & & & - & x_3 & = & -27 \end{array}$$

Finalmente, haciendo sustitución regresiva, es decir, despejando y sustituyendo variables de este último sistema de abajo hacia arriba, tenemos

$$\begin{aligned} x_3 &= 3; \\ x_2 &= \frac{-17+7(x_3)}{2} \\ &= \frac{-17+7(3)}{2} \\ &= 2; \\ x_1 &= 9 - x_2 - 2x_3 \\ &= 9-(2)-2(3) \\ &= 1. \end{aligned}$$

Así, el sistema es consistente con solución única

$$\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}. \blacktriangleleft$$

Podemos sintetizar el método del ejemplo precedente de la siguiente manera. Denotemos por R_i la i-ésima ecuación de un sistema lineal; la notación $R_i \leftrightarrow \alpha R_i + \beta R_j$ significa que la ecuación R_i se sustituye por la ecuación que se obtiene de sumar α veces la ecuación R_i con β veces la ecuación R_j. Las operaciones algebraicas que hicimos en el ejemplo anterior se resumen en el siguiente esquema.

[5] Cuando se multiplica una ecuación por un número, significa que ambos lados de la igualdad en dicha ecuación se multiplican por ese número; y cuando se suman dos ecuaciones, quiere decir que se suman miembro a miembro los correspondientes lados de la igualdad.

$$\begin{array}{rcl} x_1+x_2+2x_3 &=& 9 \\ 2x_1+4x_2-3x_3 &=& 1 \\ 3x_1+6x_2-5x_3 &=& 0 \end{array} \xleftrightarrow[R_3 \leftrightarrow -3R_1+R_3]{R_2 \leftrightarrow -2R_1+R_2} \begin{array}{rcr} x_1+x_2+2x_3 &=& 9 \\ 2x_2-7x_3 &=& -17 \\ 3x_2-11x_3 &=& -27 \end{array}$$

$$\xleftrightarrow{R_3 \leftrightarrow -3R_2+2R_3} \begin{array}{rcr} x_1+x_2+2x_3 &=& 9 \\ 2x_2-7x_3 &=& -17 \\ -x_3 &=& -3 \end{array}$$

En cada paso del proceso anterior se obtiene un sistema equivalente; es decir, con las mismas soluciones pero más sencillo, hasta que el último sistema equivalente está escalonado y se puede resolver haciendo sustitución regresiva.

Es claro que en el ejemplo 1.21 y en la discusión anterior sólo se trabajó con los coeficientes, y que de las variables x_1, x_2 y x_3 únicamente se utiliza la posición que tienen en el arreglo. Se ve entonces que para resolver un sistema lineal $A\vec{x}=\vec{b}$, basta trabajar con la matriz de coeficientes A y el término *independiente* $\vec{b}$.[6] Para ello, a continuación damos el siguiente concepto.

Definición 1.10 *Para el sistema lineal*

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

o, en forma matricial, $A\vec{x}=\vec{b}$ *con*

$$\vec{x}=\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \quad y \quad \vec{b}=\begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix},$$

se define la **matriz aumentada** *(también se le llama matriz* **ampliada***) del mismo como*

$$[A\,|\,\vec{b}\,]=\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

El lado izquierdo en la partición $[A\,|\,\vec{b}\,]$ contiene la matriz de coeficientes $[a_{ij}]$ y el lado derecho contiene los términos independientes b_i del sistema lineal. La definición anterior provee una notación muy simple para evitar, en un sistema lineal, escribir las variables y únicamente trabajar con los coeficientes. La primera fila de la matriz ampliada equivale a la ecuación $a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1$, la segunda

[6] Llamaremos término independiente en un sistema lineal $A\vec{x}=\vec{b}$, a la matriz columna $\vec{b}$ y términos independientes del mismo sistema a las respectivas componentes de este vector.

fila equivale a la ecuación $a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$, etc., y la última fila equivale a la ecuación $a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$. La línea vertical en la partición $[A \,|\, \vec{b}\,]$ únicamente sirve para hacer notoria la columna que contiene los términos independientes b_i del sistema lineal; y de hecho se puede omitir, si así se desea, cuando se conviene en que la última columna de la matriz aumentada contenga el término independiente $\vec{b}$ del sistema.

Resolveremos ahora, en el siguiente ejercicio, el ejemplo 1.21 utilizando la matriz aumentada.

▶ **Ejemplo 1.22** Para este caso, haciendo las mismas operaciones que en la discusión posterior al ejemplo 1.21, pero esta vez a los renglones de la matriz ampliada se tiene:

$$\left[\begin{array}{rrr|r} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right] \xleftrightarrow[R_3 \leftrightarrow -3R_1 + R_3]{R_2 \leftrightarrow -2R_1 + R_2} \left[\begin{array}{rrr|r} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 3 & -11 & -27 \end{array}\right]$$
$$\xleftrightarrow{R_3 \leftrightarrow -3R_2 + 2R_3} \left[\begin{array}{rrr|r} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 0 & -1 & -3 \end{array}\right]$$

y, al hacer sustitución regresiva como se hizo en ese ejemplo (cfr. pág. 18),

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}. \blacktriangleleft$$

Hasta aquí, aunque se ha utilizado en forma intuitiva el significado de sistema escalonado, no se ha precisado con exactitud. En la siguiente subsección nos abocamos a ello.

1.2.2 Matrices escalonadas y sistemas escalonados

Definición 1.11 *La matriz $A \in \mathfrak{M}_{m\times n}$ está en* **forma escalonada** *si se cumplen las siguientes dos condiciones.*

- *Las filas nulas (si existen)*[7] *están por debajo de las filas no nulas.*
- *El primer elemento distinto de cero de cada fila no nula está a la derecha del primer elemento diferente de cero de las filas precedentes.*[8]

▶ **Ejemplo 1.23** Si

$$A = \begin{bmatrix} 0 & -1 & 2 & 3 & -5 & 3 \\ 0 & 0 & -1 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} -1 & 2 & 4 & 0 & 3 \\ 0 & 1 & 2 & -3 & 4 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 2 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

A está en forma escalonada pero B no.◀

[7] Una fila es nula si todas sus entradas son ceros; una fila es no nula si por lo menos una de sus componentes es distinta de cero.

[8] En el caso que el primer elemento distinto de cero esté en la primera fila, se sobreentiende que la condición se cumple por vacuidad.

Definición 1.12 *Al primer elemento distinto de cero de cada fila no nula, de una matriz en forma escalonada, se le llama* **pivote**.

Definición 1.13 *Un sistema $H\vec{x} = \vec{c}$ está* ***escalonado*** *si la matriz ampliada $[H\,|\,\vec{c}]$ es una matriz escalonada. A las variables que correspondan a pivotes en un sistema escalonado se les llamarán* **variables ligadas** *(o principales o básicas) y a las restantes* **variables libres** *(o no básicas).*

▶ **Ejemplo 1.24** En el sistema escalonado 4×6

$$\left[\begin{array}{cccccc|c} \boxed{1} & 0 & 3 & -2 & 1 & 5 & -2 \\ 0 & 0 & \boxed{5} & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & \boxed{7} & 6 & 7 \\ 0 & 0 & 0 & 0 & 0 & \boxed{5} & 0 \end{array}\right],$$

hay pivotes en las columnas 1, 3, 5 y 6; que corresponden, respectivamente, a las variables x_1, x_3, x_5 y x_6. Así que estas variables son ligadas y x_2, x_4 son variables libres. ◀

Entonces, para resolver un sistema escalonado al hacer *sustitución regresiva*, **se despejan las variables ligadas dejándolas en función de las variables libres procediendo de abajo hacia arriba**, en el caso que el sistema tenga variables libres; en caso contrario, simplemente se despejan las variables ligadas actuando también de abajo hacia arriba.

▶ **Ejemplo 1.25** Resolver los siguientes sistemas lineales escalonados.

1. $$\left[\begin{array}{ccc|c} -5 & -1 & 3 & 3 \\ 0 & 3 & 5 & 8 \\ 0 & 0 & 2 & -4 \end{array}\right]$$

2. $$\left[\begin{array}{ccccc|c} 1 & -3 & 0 & 5 & 0 & 4 \\ 0 & 0 & 1 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

3. $$\left[\begin{array}{ccc|c} 1 & -3 & 5 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & -1 \end{array}\right]$$ ◀

Solución 1. En este caso, x_1, x_2 y x_3 son todas variables ligadas, el sistema no tiene variables libres y $x_3 = -4/2 = -2$; $x_2 = \frac{8-5x_2}{3} = 6$; $x_1 = \frac{3+x_2-3x_3}{-5} = -3$. Es decir,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3 \\ 6 \\ -2 \end{bmatrix}$$

es la única solución.

2. Para este sistema escalonado x_1, x_3 y x_5 son las variables ligadas; mientras que x_2 y x_4 son las variables libres. Entonces $x_5 = 1$, $x_3 = -7-2x_4$, $x_1 = 4+3x_2-5x_4$; lo cual indica que al dar valores concretos arbitrarios a las variables libres x_2 y x_4 se obtiene una solución. Así, el conjunto de soluciones de este sistema es infinito y está dado por:

$$\{(x_1,x_2,x_3,x_4,x_5)\,|\,x_5 = 1, x_3 = -7-2x_4, x_1 = 4+3x_2-5x_4; x_2,x_4 \in \mathbb{R}\}.$$

Una manera más compacta de expresar las soluciones es:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 4+3s-5r \\ s \\ -7-2r \\ r \\ 1 \end{bmatrix}; \quad r,s \in \mathbb{R}.$$

Al dar valores concretos a r y s se obtendrá una solución particular; por ejemplo, si $r = 0$ y $s = 0$, es fácil darse cuenta que

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ -7 \\ 0 \\ 1 \end{bmatrix}$$

resuelve el sistema de ecuaciones.

3. Para este sistema no pueden existir números reales x_1,x_2,x_3 tales que $0x_1+0x_2+0x_3 = -1$; es decir, el sistema **no tiene solución**, es inconsistente. ✔

1.2.3 Operaciones de renglón para matrices, equivalencia por filas y soluciones de sistemas escalonados

Motivados en los métodos de la subsección precedente para resolver sistemas lineales, definimos las siguientes operaciones de renglón (fila) para matrices.

Operaciones elementales de renglón para matrices

1. Intercambio de filas: $R_i \longleftrightarrow R_j$.
2. Cambio de escala: $R_i \longleftrightarrow \alpha R_i$ $(\alpha \neq 0)$.
3. Suma de filas: $R_i \longleftrightarrow \alpha R_i + \beta R_j$ $(\alpha \neq 0)$.

Las cuales significan, respectivamente:

- La fila i se intercambia con la fila j.
- La fila i se cambia por la misma fila multiplicada por α.
- La fila i se cambia por la suma de α-veces la fila i con β-veces la fila j.

Matrices equivalentes

Definición 1.14 *Sean* $A, B \in \mathfrak{M}_{m\times n}$. *B es* **equivalente por filas** *a la matriz A (o simplemente equivalente a A), si B se puede obtener de la matriz A al aplicarle una sucesión finita de operaciones elementales de renglón. Si B es equivalente a A escribiremos* $B \sim A$ *o* $B \leftrightarrow A$.

► **Ejemplo 1.26** Si

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & -3 & -1 & 0 & 1 \end{bmatrix}$$

y

$$B = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -7 & -7 & -8 & -9 \end{bmatrix},$$

$B \sim A$; pues B se obtiene de A mediante la operación de renglón

$$R_2 \longleftrightarrow -2R_1 + R_2 ◄$$

No es difícil probar el siguiente teorema.

Teorema 1.1 *Si* $A, B \in \mathfrak{M}_{m\times n}$, *entonces*

1. $A \sim A$. (Reflexividad)
2. $A \sim B \Rightarrow B \sim A$. (Simetría)
3. $A \sim B$ y $B \sim C \Rightarrow A \sim C$. (Transitividad)

Del teorema precedente inciso 2, se ve que ya no es necesario decir que B es equivalente a A, pues en tal caso simplemente podremos enunciar que A y B son equivalentes.

Al aplicar operaciones de renglón a un sistema se obtiene un sistema equivalente. Es decir:

Teorema 1.2 *Si* $[A \,|\, \vec{b}\,] \sim [H \,|\, \vec{c}\,]$, *entonces los sistemas* $A\vec{x} = \vec{b}$ *y* $H\vec{x} = \vec{c}$ *tienen las mismas soluciones.*

Es claro que siempre se pueden aplicar operaciones de fila a una matriz A, de manera adecuada, para obtener una matriz escalonada equivalente a ella. Lo cual hacemos patente en la siguiente proposición.

Teorema 1.3 *Toda matriz es equivalente por filas al menos a una matriz en forma escalonada.*

Soluciones de sistemas escalonados

Del ejemplo 1.25 (cfr. pág. 21) se conjetura el siguiente teorema, cuya demostración es sencilla y se deja como ejercicio al lector.

Teorema 1.4 *Sea un sistema* $A\vec{x} = \vec{b}$ *y supongamos que* $[H\,|\,\vec{c}]$ *es un sistema (cualquier sistema)* escalonado *equivalente; es decir,* $[A\,|\,\vec{b}] \sim [H\,|\,\vec{c}]$*, entonces*

1. $A\vec{x} = \vec{b}$ *es* **inconsistente** *si y sólo si* $[H\,|\,\vec{c}]$ *tiene una fila de ceros en el lado izquierdo y un elemento no nulo en el lado derecho de la partición (ejemplo 1.25 inciso 3).*
2. $A\vec{x} = \vec{b}$ *tiene* **solución única** *si y sólo si es consistente y* $[H\,|\,\vec{c}]$ *tiene pivote en todas las columnas en el lado izquierdo de la partición (ejemplo 1.25 inciso 1).*
3. $A\vec{x} = \vec{b}$ *tiene* **infinidad de soluciones** *si y sólo si es consistente y* $[H\,|\,\vec{c}]$ *no tiene pivote en alguna columna en el lado izquierdo de la partición (ejemplo 1.25 inciso 2).*

❍ **Nota 1.3** Es claro, del teorema anterior, que

- un sistema consistente tiene solución única cuando una forma escalonada equivalente no tiene variables libres.
- un sistema consistente tiene una infinidad de soluciones cuando una forma escalonada equivalente tiene variables libres.

1.2.4 *Método de Gauss*

El método de Gauss sirve para "llevar" una matriz a una forma escalonada equivalente aplicando operaciones de renglón. Bosquejamos el método por medio del siguiente algoritmo:

Supongamos que A es una matriz $m \times n$ no nula (si A es la matriz cero, A está en forma escalonada).

G1: Se busca una fila en A que tenga su primer elemento distinto de cero y se *intercambia* (si es necesario) con la primera fila de la matriz A; si no existe una fila de A que tenga su primer elemento no nulo, entonces se busca una fila de la matriz A que tenga el segundo elemento distinto de cero y se *intercambia* (si es necesario) con la primera fila de la matriz A; de no suceder así, se busca una fila de A que tenga el tercer elemento distinto de cero y se intercambia (si es necesario) con la primera fila de A, etc.; obteniendo finalmente una matriz $B_1 \sim A$ con un primer elemento no nulo en la primera fila que llamaremos *pivote* (en este caso de la primera fila).

Por ejemplo, si

$$A = \begin{bmatrix} 0 & -4 & -1 & 3 \\ 3 & 4 & 0 & 7 \\ 1 & 1 & 3 & 5 \end{bmatrix},$$

entonces una operación de renglón para llevar a cabo este paso puede ser $R_1 \leftrightarrow R_3$, resultando la equivalencia de matrices

$$\overbrace{\begin{bmatrix} 0 & -4 & -1 & 3 \\ 3 & 4 & 0 & 7 \\ 1 & 1 & 3 & 5 \end{bmatrix}}^{A} \sim \overbrace{\begin{bmatrix} 1 & 1 & 3 & 5 \\ 3 & 4 & 0 & 7 \\ 0 & -4 & -1 & 3 \end{bmatrix}}^{B_1}.$$

El pivote de la primera fila de la matriz B_1 es $b_{11}^1 = 1$.

G2: Con el *pivote* de la primera fila de B_1 se transforman en ceros los elementos que están por debajo de él mediante la operación *suma de filas*, obteniendo una matriz $B_2 \sim B_1 \sim A$, que tendrá todas las componentes nulas debajo del pivote de la primera fila.

Por ejemplo, con el caso particular ilustrado en el paso anterior podemos hacer ceros los elementos debajo del pivote **1** de la primera fila de la matriz B_1 mediante la operación $R_2 \leftrightarrow -3R_1 + R_2$ para obtener la matriz B_2; es decir,

$$\overbrace{\begin{bmatrix} \boxed{1} & 1 & 3 & 5 \\ 3 & 4 & 0 & 7 \\ 0 & -4 & -1 & 3 \end{bmatrix}}^{B_1} \sim \overbrace{\begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -9 & -8 \\ 0 & -4 & -1 & 3 \end{bmatrix}}^{B_2}$$

G3: Ahora se repiten los pasos G1 y G2 con la segunda fila de la matriz B_2, produciendo una matriz $B_3 \sim B_2 \sim B_1 \sim A$ cuyas componentes serán nulas debajo del pivote de su segunda fila.

Para el caso particular ilustrado, el pivote de la segunda fila de la matriz[9] B_2 es $b_{22}^2 = \mathbf{1}$. Se pueden hacer ceros los elementos debajo del mismo mediante la operación $R_3 \leftrightarrow 4R_2 + R_3$, esto es

$$\overbrace{\begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & \boxed{1} & -9 & -8 \\ 0 & -4 & -1 & 3 \end{bmatrix}}^{B_2} \sim \overbrace{\begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -9 & -8 \\ 0 & 0 & -37 & -29 \end{bmatrix}}^{B_3}$$

G4: Se repiten los pasos G1, G2 y G3 con las filas subsecuentes de las matrices equivalentes que resulten, hasta obtener una matriz H en forma escalonada de acuerdo a la definición 1.11.

Para el caso ilustrado previamente, la matriz B_3 ya está en forma escalonada; con lo que

$$\overbrace{\begin{bmatrix} 0 & -4 & -1 & 3 \\ 3 & 4 & 0 & 7 \\ 1 & 1 & 3 & 5 \end{bmatrix}}^{A} \sim \overbrace{\begin{bmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -9 & -8 \\ 0 & 0 & -37 & -29 \end{bmatrix}}^{B_3=H}$$

terminaría el proceso para este ejemplo particular.

❍ Nota 1.4

1. El lector debe tener en mente que el propósito fundamental del método de Gauss es obtener una matriz en forma escalonada equivalente a una matriz dada, mediante el uso de las operaciones elementales de renglón en cualquier combinación. Así que el algoritmo anterior sólo es una guía para este propósito. Cualquier modificación es válida siempre y cuando se empleen únicamente las operaciones de renglón para matrices y se alcance el objetivo de obtener una matriz en forma escalonada equivalente por filas a la matriz inicial.
2. A lo largo de este texto haremos uso de oraciones informales como "llevar la matriz A a forma escalonada". Este tipo de oraciones en realidad deben interpretarse como "obtener una forma

[9] El número **2** en b_{22}^2 de esta notación juega el papel de un supraíndice, haciendo referencia a la matriz B_2 y no de un exponente.

escalonada equivalente a la matriz A”, que sería la manera apropiada de expresar este tipo de instrucciones; pero, ya que esa forma escalonada equivalente se obtiene a partir de la matriz A, nos permitiremos ese tipo de frases sacrificando rigor en aras de brevedad en el lenguaje. Sin embargo, es conveniente que el lector tenga siempre presente el significado preciso de esas oraciones coloquiales.

▶ **Ejemplo 1.27** Obtener una matriz equivalente por filas a la matriz

$$A = \begin{bmatrix} 2 & -4 & 2 & -2 \\ 2 & -4 & 3 & -4 \\ 4 & -8 & 3 & -2 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

que esté en forma escalonada.[10] ◀

Solución

$$A = \begin{bmatrix} 2 & -4 & 2 & -2 \\ 2 & -4 & 3 & -4 \\ 4 & -8 & 3 & -2 \\ 0 & 0 & -1 & 2 \end{bmatrix} \xleftrightarrow{R_1 \leftrightarrow (1/2)R_1} \begin{bmatrix} \boxed{1} & -2 & 1 & -1 \\ 2 & -4 & 3 & -4 \\ 4 & -8 & 3 & -2 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

$$\xleftrightarrow[R_3 \leftrightarrow -4R_1 + R_3]{R_2 \leftrightarrow -2R_1 + R_2} \begin{bmatrix} 1 & -2 & 1 & -1 \\ 0 & 0 & \boxed{1} & -2 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

$$\xleftrightarrow[R_4 \leftrightarrow R_2 + R_4]{R_3 \leftrightarrow R_2 + R_3} \begin{bmatrix} 1 & -2 & 1 & -1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= H$$

La matriz resultante, H, está en forma escalonada y es equivalente a la matriz A. ✔

Método de Gauss para resolver sistemas lineales

▶ **Ejemplo 1.28** Resolver el siguiente sistema de ecuaciones por el método de Gauss.

$$\begin{aligned} x_1 - 2x_2 + x_3 - x_4 &= 4 \\ 2x_1 - 3x_2 + 2x_3 - 3x_4 &= -1 \\ 3x_1 - 5x_2 + 3x_3 - 4x_4 &= 3 \\ -x_1 + x_2 - x_3 + 2x_4 &= 5 \end{aligned}$$ ◀

Solución Para resolver el problema llevaremos la matriz aumentada a una forma escalonada y haremos sustitución regresiva.[11]

[10] Hemos marcado en color rojo los pivotes en cada paso para que el lector recuerde que el propósito es ir haciendo ceros, mediante las operaciones de renglón indicadas, los elementos debajo de ellos.

[11] De aquí en adelante, salvo algunas excepciones, ya no indicaremos las operaciones de renglón que se requieren para obtener una forma escalonada equivalente a una matriz, pues el objetivo es utilizar la notación matricial para auxiliarse y hacer todos los cálculos mecánica y mentalmente.

$$\left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 2 & -3 & 2 & -3 & -1 \\ 3 & -5 & 3 & -4 & 3 \\ -1 & 1 & -1 & 2 & 5 \end{array}\right] \longleftrightarrow \left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 0 & 1 & 0 & -1 & -9 \\ 0 & 1 & 0 & -1 & -9 \\ 0 & -1 & 0 & 1 & 9 \end{array}\right]$$
$$\longleftrightarrow \left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 0 & 1 & 0 & -1 & -9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Así, las variables ligadas son x_1, x_2 y las libres x_3, x_4. Y $x_2 = -9 + x_4$; $x_1 = 4 + 2x_2 - x_3 + x_4 = -14 + 3x_4 - x_3$. La solución está dada entonces por:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -14 + 3r - s \\ -9 + r \\ s \\ r \end{bmatrix}; \quad r, s \in \mathbb{R}. \quad ✔$$

Sistemas con la misma matriz de coeficientes

Es frecuente en la práctica tener que resolver sistemas con la misma matriz de coeficientes pero con distintos términos independientes; por ejemplo, los sistemas

$$\begin{array}{rcr} x - 2y + 3z & = & -2 \\ -x + 4y + 5z & = & 7 \end{array} \tag{1.14}$$

y

$$\begin{array}{rcr} r - 2s + 3t & = & 1 \\ -r + 4s + 5t & = & -4 \end{array} \tag{1.15}$$

tienen la misma matriz de coeficientes, $\begin{bmatrix} 1 & -2 & 3 \\ -1 & 4 & 5 \end{bmatrix}$, y términos independientes $\begin{bmatrix} -2 \\ 7 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ -4 \end{bmatrix}$, respectivamente. En lugar de resolverlos cada uno por separado, podemos solucionarlos simultáneamente colocando en el lado derecho de la partición de la matriz ampliada las dos columnas que contienen los dos términos independientes, llevar a forma escalonada y resolver por sustitución regresiva para la primera columna y después para la segunda:

$$\left[\begin{array}{rrr|rr} 1 & -2 & 3 & -2 & 1 \\ -1 & 4 & 5 & 7 & -4 \end{array}\right] \leftrightarrow \left[\begin{array}{rrr|rr} 1 & -2 & 3 & -2 & 1 \\ 0 & 2 & 8 & 5 & -3 \end{array}\right]. \tag{1.16}$$

Resolviendo para la primera columna tenemos $y = \frac{5}{2} - 4z$, $x = -2 + 2y - 3z = 3 - 11z$; así que

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 - 11\alpha \\ \frac{5}{2} - 4\alpha \\ \alpha \end{bmatrix},$$

$\alpha \in \mathbb{R}$, es la solución para el sistema (1.14). Resolviendo ahora para la segunda columna de (1.16) obtenemos $s = -\frac{3}{2} - 4t$, $r = 1 + 2s - 3t = -2 - 11t$; es decir,

$$\begin{bmatrix} r \\ s \\ t \end{bmatrix} = \begin{bmatrix} -2 - 11\beta \\ -\frac{3}{2} - 4\beta \\ \beta \end{bmatrix},$$

$\beta \in \mathbb{R}$, es la solución del sistema (1.15).

1.2.5 Método de Gauss-Jordan y sistemas con solución única

Definición 1.15 *Una matriz está en forma* **escalonada reducida** *si:*

1. *Está en forma escalonada.*
2. *Arriba de cada pivote las componentes (si hay) son nulas.*
3. *Todos los pivotes son unos.*

► **Ejemplo 1.29** La matriz $\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 8 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ está en forma escalonada reducida.◄

Método de Gauss-Jordan

Para llevar una matriz a forma escalonada reducida se procede de la manera siguiente:

1. Se lleva la matriz a forma escalonada mediante el método de Gauss.
2. Se hacen ceros todos los elementos arriba de cada pivote utilizando el método de Gauss de abajo hacia arriba.
3. Se convierten en unos todos los pivotes mediante la operación de renglón cambio de escala.

Empleando el método de Gauss-Jordan se puede probar el teorema que enunciamos a continuación.

Teorema 1.5 *Toda matriz es equivalente por filas a una y sólo una matriz en forma escalonada reducida.*[12]

► **Ejemplo 1.30** Obtener la forma escalonada reducida equivalente a la matriz A por el método de Gauss-Jordan si

$$A = \begin{bmatrix} -2 & -1 & 0 & 3 & -4 \\ 3 & -1 & 2 & 0 & 3 \\ 5 & -4 & 0 & -1 & 2 \end{bmatrix}.◄$$

Solución

$$\begin{bmatrix} \mathbf{-2} & -1 & 0 & 3 & -4 \\ 3 & -1 & 2 & 0 & 3 \\ 5 & -4 & 0 & -1 & 2 \end{bmatrix} \overset{(1)}{\longleftrightarrow} \begin{bmatrix} -2 & -1 & 0 & 3 & -4 \\ 0 & \mathbf{-5} & 4 & 9 & -6 \\ 0 & -13 & 0 & 13 & -16 \end{bmatrix}$$

$$\overset{(2)}{\longleftrightarrow} \begin{bmatrix} -2 & -1 & 0 & 3 & -4 \\ 0 & -5 & 4 & 9 & -6 \\ 0 & 0 & -52 & -52 & -2 \end{bmatrix}$$

[12] Compare con el teorema 1.3, página 23.

$$\overset{(3)}{\longleftrightarrow}\begin{bmatrix} -2 & -1 & 0 & 3 & -4 \\ 0 & -5 & 4 & 9 & -6 \\ 0 & 0 & -26 & -26 & -1 \end{bmatrix}$$

$$\overset{(4)}{\longleftrightarrow}\begin{bmatrix} -2 & -1 & 0 & 3 & -4 \\ 0 & -65 & 0 & 65 & -80 \\ 0 & 0 & -26 & -26 & -1 \end{bmatrix}$$

$$\overset{(5)}{\longleftrightarrow}\begin{bmatrix} -2 & -1 & 0 & 3 & -4 \\ 0 & -13 & 0 & 13 & -16 \\ 0 & 0 & -26 & -26 & -1 \end{bmatrix}$$

$$\overset{(6)}{\longleftrightarrow}\begin{bmatrix} -26 & 0 & 0 & 26 & -36 \\ 0 & -13 & 0 & 13 & -16 \\ 0 & 0 & -26 & -26 & -1 \end{bmatrix}$$

$$\overset{(7)}{\longleftrightarrow}\begin{bmatrix} 1 & 0 & 0 & -1 & 18/13 \\ 0 & 1 & 0 & -1 & 16/13 \\ 0 & 0 & 1 & 1 & 1/26 \end{bmatrix}.$$

Donde, para facilitar su comprensión, esta vez hemos indicado las operaciones de renglón en cada paso del (1) al (7), señalando los pivotes en azul más claro cuando se hacen ceros los elementos por debajo de los mismos y en rojo cuando se hacen ceros los elementos por encima de los pivotes. (1): $R_2 \leftrightarrow 3R_1 + 2R_2$, $R_3 \leftrightarrow 5R_1 + 2R_3$; (2): $R_3 \leftrightarrow -13R_2 + 5R_3$; (3): $R_3 \leftrightarrow (1/2)R_3$; (4): $R_2 \leftrightarrow 2R_3 + 13R_2$; (5): $R_2 \leftrightarrow (1/5)R_2$; (6): $R_1 \leftrightarrow 13R_1 - R_2$; (7): $R_1 \leftrightarrow (-1/26)R_1$, $R_2 \leftrightarrow (-1/13)R_2$, $R_3 \leftrightarrow (-1/26)R_3$. ✔

❍ **Nota 1.5** A diferencia de la forma escalonada reducida de una matriz, que es única, es claro que al hacer operaciones de renglón a una matriz A para obtener una matriz en forma escalonada equivalente, se pueden obtener diferentes matrices. Sin embargo, para cualquier par de matrices en forma escalonada equivalentes a la matriz A se cumple:

1. Las dos matrices tienen el mismo número de pivotes.
2. Los pivotes se encuentran en las mismas posiciones en ambas matrices; es decir, si una matriz tiene un pivote en la componente ij la otra también tiene un pivote en esta componente.

Ilustramos la nota 1.5 en el siguiente ejemplo.

► **Ejemplo 1.31** Sea $A = \begin{bmatrix} 2 & -3 & 1 & 1 & 5 \\ 4 & -1 & 0 & 1 & 2 \\ 1 & -2 & 3 & 1 & 1 \end{bmatrix}$, entonces

$$\begin{bmatrix} 2 & -3 & 1 & 1 & 5 \\ 4 & -1 & 0 & 1 & 2 \\ 1 & -2 & 3 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & -3 & 1 & 1 & 5 \\ 0 & 5 & -2 & -1 & -8 \\ 0 & 1 & -5 & -1 & 3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & -3 & 1 & 1 & 5 \\ 0 & 5 & -2 & -1 & -8 \\ 0 & 0 & 23 & 4 & -23 \end{bmatrix} = H_1$$

y

$$\begin{bmatrix} 2 & -3 & 1 & 1 & 5 \\ 4 & -1 & 0 & 1 & 2 \\ 1 & -2 & 3 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 & 1 & 1 \\ 4 & -1 & 0 & 1 & 2 \\ 2 & -3 & 1 & 1 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 & 1 & 1 \\ 0 & 7 & -12 & -3 & -2 \\ 0 & 1 & -5 & -1 & 3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -2 & 3 & 1 & 1 \\ 0 & 1 & -5 & -1 & 3 \\ 0 & 7 & -12 & -3 & -2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -2 & 3 & 1 & 1 \\ 0 & 1 & -5 & -1 & 3 \\ 0 & 0 & 23 & 4 & -23 \end{bmatrix} = H_2.$$

Así $A \sim H_1$, $A \sim H_2$; H_1 y H_2 están en forma escalonada, $H_1 \neq H_2$; ambas matrices tienen el mismo número de pivotes y se encuentran en las mismas posiciones en las dos matrices.◄

Sistemas lineales y método de Gauss-Jordan

Los sistemas lineales también se pueden resolver utilizando el método de Gauss-Jordan para llevar la matriz ampliada a forma escalonada reducida y realizar sustitución regresiva, como hacemos patente en el siguiente ejemplo.

► **Ejemplo 1.32** Resolver el siguiente sistema mediante el método de Gauss-Jordan

$$\begin{array}{rcrcrcrcr} x_1 & - & 2x_2 & + & x_3 & + & 3x_4 & = & -1 \\ -x_1 & + & x_2 & & & + & 2x_4 & = & 2 \\ 2x_1 & - & x_2 & + & x_3 & + & 5x_4 & = & 1 \end{array}$$ ◄

Solución Llevemos la matriz ampliada a la forma escalonada reducida:

$$\left[\begin{array}{rrrr|r} 1 & -2 & 1 & 3 & -1 \\ -1 & 1 & 0 & 2 & 2 \\ 2 & -1 & 1 & 5 & 1 \end{array}\right] \sim \left[\begin{array}{rrrr|r} 1 & -2 & 1 & 3 & -1 \\ 0 & -1 & 1 & 5 & 1 \\ 0 & 3 & -1 & -1 & 3 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & -2 & 1 & 3 & -1 \\ 0 & -1 & 1 & 5 & 1 \\ 0 & 0 & 2 & 14 & 6 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & -2 & 1 & 3 & -1 \\ 0 & 1 & -1 & -5 & -1 \\ 0 & 0 & 1 & 7 & 3 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & -2 & 0 & -4 & -4 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 7 & 3 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 7 & 3 \end{array}\right].$$

Al hacer sustitución regresiva tenemos $x_3 = 3 - 7x_4$, $x_2 = 2 - 2x_4$ y $x_1 = 0$; luego la solución viene dada por

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 2-2r \\ 3-7r \\ r \end{bmatrix}; \quad r \in \mathbb{R}. \quad ✔$$

Sistemas con solución única

En este breve apartado damos criterios para determinar cuándo hay solución única en un sistema utilizando la forma escalonada reducida, los cuales son fáciles de probar utilizando el teorema 1.4.

Sea $A \in \mathfrak{M}_{m\times n}$:

1. **Caso** $m > n$. Sea $A\vec{x} = \vec{b}$ un sistema lineal consistente, entonces las dos condiciones siguientes son equivalentes ((a) $\Leftrightarrow$ (b)):

   (a) El sistema $A\vec{x} = \vec{b}$ tiene solución única.

   (b) La forma escalonada reducida equivalente a A consiste de la identidad $n \times n$ seguida de $m - n$ filas nulas.

2. **Caso** $m < n$. Supongamos que el sistema $A\vec{x} = \vec{b}$ es consistente y tiene menos ecuaciones que incógnitas. Entonces tiene una infinidad de soluciones.
3. **Caso** $m = n$. $A\vec{x} = \vec{b}$ tiene solución única para todo $\vec{b}$ si y sólo si A es equivalente a la identidad; es decir, la forma escalonada reducida equivalente a A es I_n.

❍ **Nota 1.6** Para determinar que una matriz cuadrada sea equivalente a la identidad, basta revisar que al llevarla a una forma escalonada toda columna en ésta tenga pivote (por su definición, éste debe ser distinto de cero); ya que entonces, por el método de Gauss-Jordan, su forma escalonada reducida equivalente será la identidad.

1.2.6 Sistemas homogéneos

Definición 1.16 *Un sistema lineal con la forma* $A\vec{x} = \vec{0}$, *donde* $\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, *se llama* **homogéneo**.

Todo sistema homogéneo es consistente pues $\vec{x} = \vec{0}$ es solución del mismo; la llamada **solución trivial**.

De los casos 1 y 2, de criterios de solución única de la subsección precedente, deducimos el siguiente teorema.

Teorema 1.6 *Sea $A \in \mathfrak{M}_{m\times n}$. Entonces*

1. *Si $m = n$, el sistema homogéneo cuadrado $A\vec{x} = \vec{0}$ tiene solución no trivial si y sólo si A* **no** *es equivalente a la identidad.*
2. *Todo sistema homogéneo $A\vec{x} = \vec{0}$ con menos ecuaciones que incógnitas ($m < n$) tiene soluciones no triviales.*

Observemos que para resolver un sistema homogéneo $A\vec{x} = \vec{0}$, no es necesario poner ceros en la última columna de la ampliación; pues todos los términos independientes son nulos y al hacer las operaciones de renglón no se verán afectados. Así que bastará llevar a forma escalonada la matriz A y hacer sustitución regresiva recordando que los términos independientes son nulos.

► **Ejemplo 1.33** Para el sistema homogéneo

$$\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ x_1 - x_2 - 3x_3 &= 0\,, \end{aligned}$$

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & -3 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 & 1 & -1 \\ 0 & -2 & -2 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix};$$

de aquí,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2r \\ -r \\ r \end{bmatrix}; \quad r \in \mathbb{R}. \blacktriangleleft$$

1.2.7 Estructura de las soluciones

El teorema a continuación describe la estructura de las soluciones de los sistemas no homogéneos en relación con los homogéneos y, de paso, nos muestra que es posible resolver al mismo tiempo el sistema no homogéneo $A\vec{x} = \vec{b}$ y el **sistema homogéneo asociado** $A\vec{x} = \vec{0}$.

Teorema 1.7 *Sea $A \in \mathfrak{M}_{m\times n}$.*

1. *Si $\vec{h}_1$, $\vec{h}_2$ son soluciones particulares del sistema homogéneo $A\vec{x} = \vec{0}$ y $\alpha \in \mathbb{R}$, entonces*
   (a) *$\vec{h}_1 + \vec{h}_2$ es también solución.*
   (b) *$\alpha\vec{h}_1$ es solución del sistema.*
2. *Sea $\vec{p}$ una solución particular del sistema no homogéneo $A\vec{x} = \vec{b}$.*
   (a) *Si $\vec{h}$ es cualquier solución del sistema homogéneo asociado $A\vec{x} = \vec{0}$, entonces $\vec{p} + \vec{h}$ es solución del sistema no homogéneo $A\vec{x} = \vec{b}$.*
   (b) *Si $\vec{p}$ es una solución particular del no homogéneo $A\vec{x} = \vec{b}$, entonces toda solución $\vec{u}$ de este sistema tiene la forma $\vec{p} + \vec{h}$, para alguna solución $\vec{h}$ del sistema lineal homogéneo asociado $A\vec{x} = \vec{0}$.*

DEMOSTRACIÓN ■ 1. *(a)* Como $\vec{h}_1, \vec{h}_2$ son soluciones, $A\vec{h}_1 = \vec{0}$ y $A\vec{h}_2 = \vec{0}$. Luego,

$$A(\vec{h}_1 + \vec{h}_2) = A\vec{h}_1 + A\vec{h}_2 = \vec{0}.$$

Por tanto, $\vec{h}_1 + \vec{h}_2$ es también solución.

(b) $A(\alpha\vec{h}_1) = \alpha(A\vec{h}_1) = \alpha\vec{0} = \vec{0}$; lo cual nos dice que $\alpha\vec{h}$ es solución.

2. *(a)* $A(\vec{p}+\vec{h}) = A\vec{p} + A\vec{h} = \vec{b} + \vec{0} = \vec{b}$; pues $\vec{p}$ es solución de $A\vec{x} = \vec{b}$ y $\vec{h}$ es solución del homogéneo.

(b) Sea $\vec{u}$ una solución de $A\vec{x} = \vec{b}$. Entonces, si $\vec{h} = \vec{u} - \vec{p}$, $A\vec{h} = A(\vec{u} - \vec{p}) = A\vec{u} - A\vec{p} = \vec{b} - \vec{b} = \vec{0}$; y $\vec{u} = \vec{p} + \vec{h}$. ■

▶ **Ejemplo 1.34** Resolver el sistema

$$\begin{array}{rcrcrcrcr} x_1 & - & 2x_2 & + & x_3 & - & x_4 & = & 4 \\ 2x_1 & - & 3x_2 & + & 2x_3 & - & 3x_4 & = & -1 \\ 3x_1 & - & 5x_2 & + & 3x_3 & - & 4x_4 & = & 3 \\ -x_1 & + & x_2 & - & x_3 & + & 2x_4 & = & 5 \end{array}$$

y encontrar la solución del sistema homogéneo asociado. ◀

Solución

$$\left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 2 & -3 & 2 & -3 & -1 \\ 3 & -5 & 3 & -4 & 3 \\ -1 & 1 & -1 & 2 & 5 \end{array}\right] \leftrightarrow \left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 0 & 1 & 0 & -1 & -9 \\ 0 & 1 & 0 & -1 & -9 \\ 0 & -1 & 0 & 1 & 9 \end{array}\right]$$

$$\leftrightarrow \left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 0 & 1 & 0 & -1 & -9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Así, la solución del sistema es

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3r - s - 14 \\ r - 9 \\ s \\ r \end{bmatrix} = \begin{bmatrix} -14 \\ -9 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 3r - s \\ r \\ s \\ r \end{bmatrix}; \quad r, s \in \mathbb{R}.$$

De donde observamos que para este caso

$$\begin{bmatrix} -14 \\ -9 \\ 0 \\ 0 \end{bmatrix} \text{ y } \begin{bmatrix} 3r - s \\ r \\ s \\ r \end{bmatrix}; \quad r, s \in \mathbb{R}$$

son, respectivamente, una solución particular del no homogéneo y la solución del homogéneo asociado. ✔

1.2.8 Sistemas lineales con números complejos

Nuevamente, como en la sección 1.1.5 (cfr. pág. 12), haremos uso de números complejos en álgebra lineal; esta vez en sistemas lineales. Para resolver sistemas lineales con coeficientes complejos simplemente aplicaremos las mismas técnicas que hemos desarrollado en esta sección, pero realizando las operaciones con números complejos. Todas las propiedades y teoremas relativos a sistemas lineales sobre el campo de los números reales de esta sección siguen siendo válidos cuando los coeficientes y términos independientes son números complejos.[13]

► **Ejemplo 1.35** Resolvamos el sistema

$$\begin{array}{rcrcrcr} x_1 & + & (1+i)x_2 & - & ix_3 & = & 2i \\ ix_1 & - & 2ix_2 & + & x_3 & = & 1-2i \\ x_1 & + & (1-i)x_2 & + & 2ix_3 & = & 1+i. \end{array}$$

Para ello aplicamos el método de Gauss empleando el álgebra de números complejos en las operaciones.[14]

$$\left[\begin{array}{ccc|c} 1 & 1+i & -i & 2i \\ i & -2i & 1 & 1-2i \\ 1 & 1-i & 2i & 1+i \end{array}\right] \overset{(1)}{\longleftrightarrow} \left[\begin{array}{ccc|c} 1 & 1+i & -i & 2i \\ 0 & 1-3i & 0 & 3-2i \\ 0 & -2i & 3i & 1-i \end{array}\right]$$

$$\overset{(2)}{\longleftrightarrow} \left[\begin{array}{ccc|c} 1 & 1+i & -i & 2i \\ 0 & -2i & 3i & 1-i \\ 0 & 1-3i & 0 & 3-2i \end{array}\right]$$

$$\overset{(3)}{\longleftrightarrow} \left[\begin{array}{ccc|c} 1 & 1+i & -i & 2i \\ 0 & 1 & -3/2 & 1/2+1/2i \\ 0 & 1-3i & 0 & 3-2i \end{array}\right]$$

$$\overset{(4)}{\longleftrightarrow} \left[\begin{array}{ccc|c} 1 & 1+i & -i & 2i \\ 0 & 1 & -3/2 & 1/2+1/2i \\ 0 & 0 & 3/2-9i/2 & 1-i \end{array}\right].$$

De donde, al hacer sustitución regresiva,

$$\begin{aligned} x_3 &= \frac{1-i}{\frac{3}{2}-\frac{9}{2}i} \\ &= \frac{4}{15}+\frac{2}{15}i, \end{aligned}$$

$$\begin{aligned} x_2 &= \frac{1}{2}+\frac{1}{2}i+\frac{3}{2}x_3 \\ &= \frac{1}{2}+\frac{1}{2}i+\frac{2}{5}+\frac{1}{5}i \\ &= \frac{9}{10}+\frac{7}{10}i, \end{aligned}$$

y

$$\begin{aligned} x_1 &= 2i-(1+i)x_2+ix_3 \\ &= 2i-(1+i)(\tfrac{9}{10}+\tfrac{7}{10}i)+i(\tfrac{4}{15}+\tfrac{2}{15}i) \\ &= -\tfrac{1}{3}+\tfrac{2}{3}i. \blacktriangleleft \end{aligned}$$

[13] Cfr. B.1 del apéndice B y la subsección 1.1.5.

[14] (1): $R_2 \leftrightarrow -iR_1+R_2$, $R_3 \leftrightarrow -R_1+R_2$; (2): $R_2 \leftrightarrow R_3$; (3): $R_2 \leftrightarrow (-1/2i)R_2$; (4): $R_3 \leftrightarrow -(1-3i)R_2+R_3$.

▶ **Ejemplo 1.36** Resolver el sistema

$$\begin{aligned} x_1 + ix_2 - x_3 &= 5 \\ x_1 - 2ix_2 + 5x_3 &= 5 + 6i. \end{aligned}$$ ◀

Solución

$$\left[\begin{array}{ccc|c} 1 & i & -1 & 5 \\ 1 & -2i & 5 & 5+6i \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & i & -1 & 5 \\ 0 & -3i & 6 & 6i \end{array}\right]$$

y al hacer sustitución regresiva

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5+2i-\alpha \\ -2-2i\alpha \\ \alpha \end{bmatrix}; \quad \alpha \in \mathbb{C}.$$

Finalmente, dado que

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5+2i \\ -2 \\ 0 \end{bmatrix} + \begin{bmatrix} -\alpha \\ -2i\alpha \\ \alpha \end{bmatrix},$$

se desprende que la solución del sistema homogéneo asociado está dada por

$$\begin{bmatrix} -\alpha \\ -2i\alpha \\ \alpha \end{bmatrix}, \quad \alpha \in \mathbb{C}.$$ ✔

1.3 Ejercicios resueltos y ejercicios propuestos

1.3.1 Ejercicios resueltos

Matrices

1 Determinar los valores de a y b, si es que existen, para que las matrices $A = \begin{bmatrix} 1 & -2 & a \\ 3 & 2 & 2 \\ 4 & -5 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & -2 & 2b \\ 3 & a & 2 \\ 4 & -5 & 1 \end{bmatrix}$ sean iguales.

Solución Las matrices A y B serán iguales si $a = 2b$ y $a = 2$, y esto sucede si $a = 2$ y $b = 1$. ✔

2 Calcular

(a) $\begin{bmatrix} -1 & 0 & 1 \\ 2 & 3 & -4 \end{bmatrix} + \begin{bmatrix} 2 & -3 & 1 \\ -5 & -3 & 8 \end{bmatrix}$.

(b) $\begin{bmatrix} -1 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 3 & -1 \end{bmatrix} + \begin{bmatrix} -2 & 3 & -1 \\ 2 & 6 & -5 \end{bmatrix}$.

(c) $-4\begin{bmatrix} -2 & 3 & -1 & 1 \\ 0 & -1 & 1 & 2 \end{bmatrix}$.

Solución

(a) $\begin{bmatrix} -1 & 0 & 1 \\ 2 & 3 & -4 \end{bmatrix} + \begin{bmatrix} 2 & -3 & 1 \\ -5 & -3 & 8 \end{bmatrix} = \begin{bmatrix} -1+2 & 0-3 & 1+1 \\ 2-5 & 3-3 & -4+8 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 0 & 4 \end{bmatrix}$.

(b) La suma de estas matrices no está definida porque tienen distintos tamaños; la primera tiene tamaño 3×3 mientras que la segunda tiene tamaño 2×3.

(c)
$$-4\begin{bmatrix}-2 & 3 & -1 & 1\\ 0 & -1 & 1 & 2\end{bmatrix} = \begin{bmatrix}(-4)(-2) & (-4)(3) & (-4)(-1) & (-4)(1)\\ (-4)(0) & (-4)(-1) & (-4)(1) & (-4)(2)\end{bmatrix}$$
$$= \begin{bmatrix}8 & -12 & 4 & -4\\ 0 & 4 & -4 & -8\end{bmatrix} \quad ✔$$

3 Calcular

(a) $\begin{bmatrix}-1 & 2 & 0\\ 1 & 2 & 1\\ -2 & 2 & -3\\ 1 & -2 & 4\end{bmatrix}\begin{bmatrix}2 & -3\\ 1 & 2\\ -3 & 2\end{bmatrix}$. *(b)* $\begin{bmatrix}2 & -3\\ 1 & 2\\ -3 & 2\end{bmatrix}\begin{bmatrix}-1 & 2 & 0\\ 1 & 2 & 1\\ -2 & 2 & -3\\ 1 & -2 & 4\end{bmatrix}$.

Solución Recordemos que el producto de dos matrices, $A\in\mathfrak{M}_{m\times n}$ y $B\in\mathfrak{M}_{r\times s}$, está definido sólo si $n=r$; en tal caso la matriz producto, AB, tiene entonces tamaño $m\times s$. Una forma de recordar esto es que en el "producto" $(m\times n)(r\times s)$ los "medios" deben ser iguales; esto es, $n=r$; y la matriz producto tiene tamaño el producto de los "extremos": $m\times s$. Finalmente, si el producto está definido, la i-ésima fila del mismo se obtiene multiplicando la fila F_i de la matriz A con cada una de las columnas C_j de la matriz B como se indica en (1.1) de la página 5.

(a) En este caso se tiene $(4\times 3)\,(3\times 2)$, los "medios" son iguales así que el producto está definido y el tamaño de la matriz producto es (4×2), el producto de los "extremos".

$$\begin{bmatrix}-1 & 2 & 0\\ 1 & 2 & 1\\ -2 & 2 & -3\\ 1 & -2 & 4\end{bmatrix}\begin{bmatrix}2 & -3\\ 1 & 2\\ -3 & 2\end{bmatrix} = \begin{bmatrix}F_1C_1 & F_1C_2\\ F_2C_1 & F_2C_1\\ F_3C_1 & F_3C_1\\ F_3C_1 & F_3C_1\end{bmatrix}$$
$$= \begin{bmatrix}-2+2+0 & 3+4+0\\ 2+2-3 & -3+4+2\\ -4+2+9 & 6+4-6\\ 2-2-12 & -3-4+8\end{bmatrix}$$
$$= \begin{bmatrix}0 & 7\\ 1 & 3\\ 7 & 4\\ -12 & 1\end{bmatrix}.$$

(b) En este caso se tiene $(3\times 2)(4\times 3)$, que tiene "medios" distintos, así que el producto no está definido. ✔

En los ejercicios 4 a 13 sean

$$A=\begin{bmatrix}-5 & -1 & 3\\ -3 & 7 & 0\end{bmatrix},\quad B=\begin{bmatrix}-2 & 1 & 4\\ 1 & 1 & -1\end{bmatrix},$$
$$C=\begin{bmatrix}4 & -1\\ 2 & 3\\ -1 & 3\end{bmatrix} \quad\text{y}\quad D=\begin{bmatrix}-1 & 1\\ 2 & 3\\ -2 & 5\end{bmatrix}.$$

Efectuar la operación indicada si es que está definida.

4 $4A$.

Solución $4A = 4\begin{bmatrix} -5 & -1 & 3 \\ -3 & 7 & 0 \end{bmatrix} = \begin{bmatrix} (4)(-5) & (4)(-1) & (4)(3) \\ (4)(-3) & (4)(7) & (4)(0) \end{bmatrix} = \begin{bmatrix} -20 & -4 & 12 \\ -12 & 28 & 0 \end{bmatrix}$. ✔

5 $-5A+2B$.

Solución
$$\begin{aligned} -5A+2B &= -5\begin{bmatrix} -5 & -1 & 3 \\ -3 & 7 & 0 \end{bmatrix} + 2\begin{bmatrix} -2 & 1 & 4 \\ 1 & 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 25 & 5 & -15 \\ 15 & -35 & 0 \end{bmatrix} + \begin{bmatrix} -4 & 2 & 8 \\ 2 & 2 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 21 & 7 & -7 \\ 17 & -33 & -2 \end{bmatrix}. \end{aligned}$$
✔

6 $3C-4D$.

Solución
$$\begin{aligned} 3C-4D &= 3\begin{bmatrix} 4 & -1 \\ 2 & 3 \\ -1 & 3 \end{bmatrix} - 4\begin{bmatrix} -1 & 1 \\ 2 & 3 \\ -2 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 12 & -3 \\ 6 & 9 \\ -3 & 9 \end{bmatrix} + \begin{bmatrix} 4 & -4 \\ -8 & -12 \\ 8 & -20 \end{bmatrix} \\ &= \begin{bmatrix} 16 & -7 \\ -2 & -3 \\ 5 & -11 \end{bmatrix}. \end{aligned}$$
✔

7 AC.

Solución
$$\begin{aligned} AC &= \begin{bmatrix} -5 & -1 & 3 \\ -3 & 7 & 0 \end{bmatrix}\begin{bmatrix} 4 & -1 \\ 2 & 3 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} F_1C_1 & F_1C_2 \\ F_2C_1 & F_2C_2 \end{bmatrix} \\ &= \begin{bmatrix} (-5)(4)+(-1)(2)+(3)(-1) & (-5)(-1)+(-1)(3)+(3)(3) \\ (-3)(4)+(7)(2)+(0)(-1) & (-3)(-1)+(7)(3)+(0)(3) \end{bmatrix} \\ &= \begin{bmatrix} -25 & 11 \\ 2 & 24 \end{bmatrix}. \end{aligned}$$
✔

8 $(3B)(5C)$.

Solución
$$\begin{aligned} (3B)(5C) &= \left(3\begin{bmatrix} -2 & 1 & 4 \\ 1 & 1 & -1 \end{bmatrix}\right)\left(5\begin{bmatrix} 4 & -1 \\ 2 & 3 \\ -1 & 3 \end{bmatrix}\right) \\ &= 15\begin{bmatrix} -2 & 1 & 4 \\ 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 4 & -1 \\ 2 & 3 \\ -1 & 3 \end{bmatrix} \\ &= 15\begin{bmatrix} -10 & 17 \\ 7 & -1 \end{bmatrix} = \begin{bmatrix} -150 & 255 \\ 105 & -15 \end{bmatrix}. \end{aligned}$$
✔

9 CD.

Solución El producto CD no está definido porque la matriz C tiene tamaño 2×3 y la matriz D tiene tamaño 2×3; el número de columnas de C (tres) es distinto del número de filas de D (dos). ✔

10 A^tA.

Solución

$$A^tA = \begin{bmatrix} -5 & -1 & 3 \\ -3 & 7 & 0 \end{bmatrix}^t \begin{bmatrix} -5 & -1 & 3 \\ -3 & 7 & 0 \end{bmatrix} = \begin{bmatrix} -5 & -3 \\ -1 & 7 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} -5 & -1 & 3 \\ -3 & 7 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 34 & -16 & -15 \\ -16 & 50 & -3 \\ -15 & -3 & 9 \end{bmatrix}. \checkmark$$

11 $(2B-3A)D$.

Solución

$$(2B-3A)D = \left(2\begin{bmatrix} -2 & 1 & 4 \\ 1 & 1 & -1 \end{bmatrix} - 3\begin{bmatrix} -5 & -1 & 3 \\ -3 & 7 & 0 \end{bmatrix}\right) \begin{bmatrix} -1 & 1 \\ 2 & 3 \\ -2 & 5 \end{bmatrix}$$

$$= \left(\begin{bmatrix} -4 & 2 & 8 \\ 2 & 2 & -2 \end{bmatrix} + \begin{bmatrix} 15 & 3 & -9 \\ 9 & -21 & 0 \end{bmatrix}\right) \begin{bmatrix} -1 & 1 \\ 2 & 3 \\ -2 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 11 & 5 & -1 \\ 11 & -19 & -2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & 3 \\ -2 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 21 \\ -45 & -56 \end{bmatrix}. \checkmark$$

12 $(AD)B$.

Solución

$$(AD)B = \left(\begin{bmatrix} -5 & -1 & 3 \\ -3 & 7 & 0 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & 3 \\ -2 & 5 \end{bmatrix}\right) \begin{bmatrix} -2 & 1 & 4 \\ 1 & 1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} -3 & 7 \\ 17 & 18 \end{bmatrix} \begin{bmatrix} -2 & 1 & 4 \\ 1 & 1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 13 & 4 & -19 \\ -16 & 35 & 50 \end{bmatrix}. \checkmark$$

13 CC^t.

Solución

$$CC^t = \begin{bmatrix} 4 & -1 \\ 2 & 3 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ 2 & 3 \\ -1 & 3 \end{bmatrix}^t = \begin{bmatrix} 4 & -1 \\ 2 & 3 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 4 & 2 & -1 \\ -1 & 3 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 17 & 5 & -7 \\ 5 & 13 & 7 \\ -7 & 7 & 10 \end{bmatrix}. \checkmark$$

14 Sean $A = \begin{bmatrix} -1 & 2 & 3 \end{bmatrix}$ y $\vec{b} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$. Calcular:

(a) $A\vec{b}$.

(b) $\vec{b}A$.

Solución *(a)* $A\vec{b} = \begin{bmatrix} -1 & 2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = 9.$

(b)

$$\vec{b}A = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \begin{bmatrix} -1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} (-1)(-1) & (-1)(2) & (-1)(3) \\ (1)(-1) & (1)(2) & (1)(3) \\ (2)(-1) & (2)(2) & (2)(3) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -2 & -3 \\ -1 & 2 & 3 \\ -2 & 4 & 6 \end{bmatrix}. \checkmark$$

15 Sea A una matriz cuadrada; se define $A^2 = AA$ y en general $A^n = \underbrace{AA \cdots A}_{n \text{ factores}}$. Calcular A^2 si $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

Solución $A^2 = AA = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$. ✔

16 Demostrar que la suma de matrices es una operación conmutativa; esto es, si $A, B \in \mathfrak{M}_{m\times n}$, entonces $A + B = B + A$.

DEMOSTRACIÓN ■ Si $A = [a_{ij}]$ y $B = [b_{ij}]$, entonces $A + B = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = B + A$ (porque la suma de números reales es una operación conmutativa). ■

17 Probar que si X es una matriz fila con n componentes y Y es una matriz columna con igual número de componentes, entonces $XY = Y^t X^t$.

DEMOSTRACIÓN ■ Sean $X = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$ y $Y = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$. Entonces $XY = \sum_{i=1}^n a_i b_i$ y

$$Y^t X^t = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \sum_{i=1}^n b_i a_i$$

$$= \sum_{i=1}^n a_i b_i = XY. \blacksquare$$

18 Demostrar la propiedad 3*(b)* de la sección 1.1.4; es decir, que $(AB)^t = B^t A^t \;\; \forall A \in \mathfrak{M}_{m\times n}, \forall B \in \mathfrak{M}_{n\times p}$.

DEMOSTRACIÓN ■ Sean A_i, $i = 1, \ldots, m$, las filas de la matriz A; B_j, $j = 1, \ldots, p$, las columnas de la matriz B. Entonces

$$(AB)^t = \begin{bmatrix} A_1B_1 & A_1B_2 & \cdots & A_1B_p \\ A_2B_1 & A_2B_2 & \cdots & A_2B_p \\ \vdots & & & \\ A_mB_1 & A_mB_2 & \cdots & A_mB_p \end{bmatrix}^t$$

$$= \begin{bmatrix} A_1B_1 & A_2B_1 & \cdots & A_mB_1 \\ A_1B_2 & A_2B_2 & \cdots & A_mB_2 \\ \vdots & \vdots & \ddots & \vdots \\ A_1B_p & A_2B_p & \cdots & A_mB_p \end{bmatrix}$$

y

$$\begin{aligned} B^tA^t &= \begin{bmatrix} B_1^t \\ B_2^t \\ \vdots \\ B_p^t \end{bmatrix} \begin{bmatrix} A_1^t & A_2^t & \cdots & A_m^t \end{bmatrix} \\ &= \begin{bmatrix} B_1^tA_1^t & B_1^tA_2^t & \cdots & B_1^tA_m^t \\ B_2^tA_1^t & B_2^tA_2^t & \cdots & B_2^tA_m^t \\ \vdots & \vdots & \ddots & \vdots \\ B_p^tA_1^t & B_p^tA_2^t & \cdots & B_p^tA_m^t \end{bmatrix} \\ &\overset{\text{(ejercicio anterior)}}{=} \begin{bmatrix} A_1B_1 & A_2B_1 & \cdots & A_mB_1 \\ A_1B_2 & A_2B_2 & \cdots & A_mB_2 \\ \vdots & \vdots & \ddots & \vdots \\ A_1B_p & A_2B_p & \cdots & A_mB_p \end{bmatrix}. \end{aligned}$$

Con lo que se tiene $(AB)^t = B^tA^t$. ■

19 Suponer que A y B son matrices cuadradas del mismo orden, ¿es cierto que $(A+B)^2 = A^2+2AB+B^2$ para todo par de matrices A y B?

Solución Si $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, entonces $(A+B)^2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ y

$$A^2+2AB+B^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

Sin embargo, si $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$, entonces $(A+B)^2 = \begin{bmatrix} 0 & -1 \\ 2 & 2 \end{bmatrix}^2 = \begin{bmatrix} -2 & -2 \\ 4 & 2 \end{bmatrix}$, pero

$$A^2+2AB+B^2 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^2 + 2\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}^2 = \begin{bmatrix} -3 & -4 \\ 2 & 3 \end{bmatrix}.$$

Es decir, en general $A^2+2AB+B^2 \neq (A+B)^2$. ✔

20 ¿Qué condición deben cumplir dos matrices cuadradas del mismo orden, A y B, para que $(A+B)^2 = A^2+2AB+B^2$?

Solución Del ejercicio precedente se puede inferir que la condición que se requiere es que las matrices conmuten; es decir, que $AB = BA$. En efecto, si éste es el caso, entonces

$$\begin{aligned} (A+B)^2 &= (A+B)(A+B) = A(A+B)+B(A+B) \\ &= A^2+AB+BA+B^2 = A^2+AB+AB+B^2 \\ &= A^2+2AB+B^2. \end{aligned}$$ ✔

21 Determinar, si es que existe, el valor que debe tener x para que las matrices $A = \begin{bmatrix} -3 & 2 \\ 1 & 4 \end{bmatrix}$ y $B = \begin{bmatrix} 3 & -4 \\ -2 & x \end{bmatrix}$ conmuten; esto es, para que $AB = BA$.

Solución Para que $AB = BA$ se debe tener

$$\begin{bmatrix} -3 & 2 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 3 & -4 \\ -2 & x \end{bmatrix} = \begin{bmatrix} 3 & -4 \\ -2 & x \end{bmatrix}\begin{bmatrix} -3 & 2 \\ 1 & 4 \end{bmatrix};$$

es decir,

$$\begin{bmatrix} -13 & 12+2x \\ -5 & -4+4x \end{bmatrix} = \begin{bmatrix} -13 & -10 \\ 6+x & -4+4x \end{bmatrix};$$

por tanto, se debe cumplir

$$\begin{aligned} 12+2x &= -10 \quad \text{y} \\ 6+x &= -5. \end{aligned}$$

Lo cual sucede únicamente si $x = -11$. ✔

22 ¿Qué condición debe cumplir una matriz cuadrada $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ para que sea simétrica?

Solución Para que A sea simétrica se debe tener $A = A^t$; esto es,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

y esto sucede si y sólo si $b = c$. ✔

23 Demostrar que si $A \in \mathfrak{M}_{m\times n}$, entonces la matriz $B = AA^t$ es simétrica.

DEMOSTRACIÓN ■ $B^t = (AA^t)^t = (A^t)^t A^t = AA^t = B$. Por tanto, B es simétrica (cfr. propiedades de la pág. 10). ■

En los ejercicios 24 a 26 sean

$$A = \begin{bmatrix} -i & 1+i & 2 \\ 1-i & 2i & 4+i \end{bmatrix}, \quad B = \begin{bmatrix} -1+i & 3i \\ -5+3i & 8 \\ 7-2i & 9i \end{bmatrix}, \quad C = \begin{bmatrix} -1 & 3+2i & -5i \\ 3 & -4i & 0 \\ 2 & -1 & 1 \end{bmatrix} \quad \text{y} \quad \vec{u} = \begin{bmatrix} -1 \\ 2-2i \\ 0 \end{bmatrix}.$$

24 Calcular

(a) $2iA$,

(b) $A - 3B^t$,

(c) $-3iB^t + 2A$.

Solución (a)

$$\begin{aligned} 2iA &= 2i\begin{bmatrix} -i & 1+i & 2 \\ 1-i & 2i & 4+i \end{bmatrix} \\ &= \begin{bmatrix} (2i)(-i) & (2i)(1+i) & (2i)(2) \\ (2i)(1-i) & (2i)(2i) & (2i)(4+i) \end{bmatrix} \\ &= \begin{bmatrix} 2 & -2+2i & 4i \\ 2+2i & -4 & -2+8i \end{bmatrix}. \end{aligned}$$

(b)
$$A-3B^t = \begin{bmatrix} -i & 1+i & 2 \\ 1-i & 2i & 4+i \end{bmatrix} - 3\begin{bmatrix} -1+i & 3i \\ -5+3i & 8 \\ 7-2i & 9i \end{bmatrix}^t$$
$$= \begin{bmatrix} -i & 1+i & 2 \\ 1-i & 2i & 4+i \end{bmatrix}$$
$$-3\begin{bmatrix} -1+i & -5+3i & 7-2i \\ 3i & 8 & 9i \end{bmatrix}$$
$$= \begin{bmatrix} -i & 1+i & 2 \\ 1-i & 2i & 4+i \end{bmatrix}$$
$$+\begin{bmatrix} 3-3i & 15-9i & -21+6i \\ -9i & -24 & -27i \end{bmatrix}$$
$$= \begin{bmatrix} 3-4i & 16-8i & -19+6i \\ 1-10i & -24+2i & 4-26i \end{bmatrix}.$$

(c)
$$-3iB^t+2A = -3i\begin{bmatrix} -1+i & 3i \\ -5+3i & 8 \\ 7-2i & 9i \end{bmatrix}^t + 2\begin{bmatrix} -i & 1+i & 2 \\ 1-i & 2i & 4+i \end{bmatrix}$$
$$= -3i\begin{bmatrix} -1+i & -5+3i & 7-2i \\ 3i & 8 & 9i \end{bmatrix} + 2\begin{bmatrix} -i & 1+i & 2 \\ 1-i & 2i & 4+i \end{bmatrix}$$
$$= \begin{bmatrix} 3+3i & 9+15i & -6-21i \\ 9 & -24i & 27 \end{bmatrix} + \begin{bmatrix} -2i & 2+2i & 4 \\ 2-2i & 4i & 8+2i \end{bmatrix}$$
$$= \begin{bmatrix} 3+i & 11+17i & -2-21i \\ 11-2i & -20i & 35+2i \end{bmatrix}. \checkmark$$

25 Encontrar

(a) $A\vec{u}$,

(b) $B\vec{u}$,

(c) $C\vec{u}$.

Solución *(a)*
$$A\vec{u} = \begin{bmatrix} -i & 1+i & 2 \\ 1-i & 2i & 4+i \end{bmatrix}\begin{bmatrix} -1 \\ 2-2i \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} (-i)(-1)+(1+i)(2-2i)+(2)(0) \\ (1-i)(-1)+(2i)(2-2i)+(4+i)(0) \end{bmatrix}$$
$$= \begin{bmatrix} 4+i \\ 3+5i \end{bmatrix}.$$

(b) La matriz B tiene tamaño 3×2 y la matriz $\vec{u}$ tiene tamaño 3×1, así que el producto $B\vec{u}$ no está definido; pues el número de columnas de B es distinto al número de filas de A. (Los "medios" en $(3\times 2)(3\times 1)$ son distintos.)

(c)
$$C\vec{u} = \begin{bmatrix} -1 & 3+2i & -5i \\ 3 & -4i & 0 \\ 2 & -1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 2-2i \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} (-1)(-1)+(3+2i)(2-2i)+(-5i)(0) \\ (3)(-1)+(-4i)(2-2i)+(0)(0) \\ (2)(-1)+(-1)(2-2i)+(1)(0) \end{bmatrix} = \begin{bmatrix} 11-2i \\ -11-8i \\ -4+2i \end{bmatrix}. \checkmark$$

26 Hallar

(a) AB, *(b)* AC, *(c)* BC.

Solución *(a)* De acuerdo con (1.1) (cfr. pág. 5)

$$\begin{aligned} AB &= \begin{bmatrix} -i & 1+i & 2 \\ 1-i & 2i & 4+i \end{bmatrix} \begin{bmatrix} -1+i & 3i \\ -5+3i & 8 \\ 7-2i & 9i \end{bmatrix} \\ &= \begin{bmatrix} F_1C_1 & F_1C_2 \\ F_2C_1 & F_2C_2 \end{bmatrix} \\ &= \begin{bmatrix} 7-5i & 11+26i \\ 24-9i & -6+55i \end{bmatrix}. \end{aligned}$$

(b) Por (1.1)

$$\begin{aligned} AC &= \begin{bmatrix} -i & 1+i & 2 \\ 1-i & 2i & 4+i \end{bmatrix} \begin{bmatrix} -1 & 3+2i & -5i \\ 3 & -4i & 0 \\ 2 & -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} F_1C_1 & F_1C_2 & F_1C_3 \\ F_2C_1 & F_2C_2 & F_2C_3 \end{bmatrix} \\ &= \begin{bmatrix} 7+4i & 4-7i & -3 \\ 7+9i & 9-2i & -1-4i \end{bmatrix}. \end{aligned}$$

(c) El producto no está definido; pues en $(3\times 2)(3\times 3)$ los “medios” son distintos (el número de columnas de B es distinto al número de filas de C). ✔

Sistemas lineales

27 Indicar si el sistema dado es lineal o no.

(a) $$\begin{aligned} x_1 - 3x_2 + x_3 &= -3 \\ x_1 - 4x_2 - 3x_3^2 &= 5 \end{aligned}$$

(b) $$\begin{aligned} x_1 - \cos(x_2)x_2 &= 0 \\ x_1 + x_2 &= 1 \end{aligned}$$

(c) $$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ x_1 - x_2 + x_3 &= 2 \end{aligned}$$

Solución *(a)* El sistema no es lineal porque no tiene la forma (1.6) (cfr. pág. 15) al estar la variable x_3 en la segunda ecuación elevada al cuadrado.

(b) El sistema no es lineal porque no tiene la forma (1.6) pues en la primera ecuación el segundo término contiene la evaluación de $\cos(x_2)$.

(c) El sistema sí es lineal pues tiene la forma (1.6); con $a_{11} = a_{12} = a_{13} = 1$, $a_{21} = a_{23} = 1$, $a_{22} = -1$, $b_1 = 3$ y $b_2 = 2$ (un sistema de dos ecuaciones con tres incógnitas). ✔

28 Determinar si $\vec{u} = (-1, 2, 1, -2)$ es solución del sistema

$$\begin{aligned} -8x_1 + 5x_2 - 3x_3 - 3x_4 &= 21 \\ 7x_1 + 5x_2 + 7x_3 + 6x_4 &= -2 \\ 4x_1 + 3x_2 + 7x_3 - 5x_4 &= 19 \end{aligned}$$

Solución Al sustituir $\vec{u}$ en la primera, segunda y tercera ecuación se obtiene:

$$\begin{aligned} -8(-1) + 5(2) - 3(1) - 3(-2) &= 21 \\ 7(-1) + 5(2) + 7(1) + 6(-2) &= -2 \\ 4(-1) + 3(2) + 7(1) - 5(-2) &= 19 \end{aligned}$$

por lo que $\vec{u}$ sí es solución del sistema. ✔

29 Determinar si $\vec{v} = (-3,0,0,1)$ es solución del sistema del ejemplo anterior.

Solución Al sustituir $\vec{v}$ en la primera y segunda ecuación se obtiene:

$$\begin{aligned} -8(-3)+5(0)-3(0)-3(1) &= 21 \\ 7(-3)+5(0)+7(0)+6(1) &= -15. \end{aligned}$$

Puesto que $-15 \neq -2$, $\vec{v}$ no es solución del sistema. ✔

30 Escribir el sistema del ejercicio 28 en la forma matricial $A\vec{x} = \vec{b}$ y resolver los ejercicios 28 y 29 mediante multiplicación de matrices.

Solución La matriz de coeficientes (cfr. pág. 15) de este sistema es $A = \begin{bmatrix} -8 & 5 & -3 & -3 \\ 7 & 5 & 7 & 6 \\ 4 & 3 & 7 & -5 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ y el término independiente es $\begin{bmatrix} 21 \\ -2 \\ 19 \end{bmatrix}$; así que el sistema en forma matricial está dado por

$$\begin{bmatrix} -8 & 5 & -3 & -3 \\ 7 & 5 & 7 & 6 \\ 4 & 3 & 7 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 21 \\ -2 \\ 19 \end{bmatrix}.$$

Puesto que

$$\begin{bmatrix} -8 & 5 & -3 & -3 \\ 7 & 5 & 7 & 6 \\ 4 & 3 & 7 & -5 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 21 \\ -2 \\ 19 \end{bmatrix}$$

y

$$\begin{bmatrix} -8 & 5 & -3 & -3 \\ 7 & 5 & 7 & 6 \\ 4 & 3 & 7 & -5 \end{bmatrix} \begin{bmatrix} -3 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 21 \\ -15 \\ -17 \end{bmatrix} \neq \vec{b},$$

$\vec{u} = (-1,2,1,-2)$ es solución del sistema $A\vec{x} = \vec{b}$, mientras que $\vec{v} = (-3,0,0,1)$ no es solución del sistema $A\vec{x} = \vec{b}$. ✔

31 Indicar las matrices que están en forma escalonada. Para las que estén en forma escalonada, determinar los pivotes de cada fila, y para las que no están en forma escalonada, mencionar la propiedad que no cumplen para este efecto.

$$(a) \begin{bmatrix} -1 & 2 & 1 & 5 & 0 \\ 0 & 2 & -3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad (b) \begin{bmatrix} 0 & 1 & -3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (c) \begin{bmatrix} 1 & 2 & -2 & 7 & 9 \\ 0 & 0 & 1 & -1 & 3 \\ 0 & 1 & -1 & 0 & 1 \end{bmatrix} \quad \text{y} \quad (d) \begin{bmatrix} -7 & 2 & 3 & -5 & 9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & -3 & 1 & 1 \end{bmatrix}.$$

Solución *(a)* La matriz está en forma escalonada; los pivotes para la primera, segunda y tercera fila son, respectivamente, -1, 2 y 1.

(b) La matriz está en forma escalonada; el pivote para la primera fila es 1 y para la segunda es 2.

(c) La matriz no está en forma escalonada; el primer elemento distinto de cero de la tercera fila no está a la derecha del primer elemento distinto de cero de la segunda fila.

(d) La fila nula en esta matriz no está por debajo de las filas no nulas; por lo que no está en forma escalonada. ✔

Para cada uno de los sistemas de los ejercicios 32 y 33:

(a) Escribir la matriz aumentada correspondiente.

(b) Utilizar el inciso *(a)* para determinar si el sistema está en forma escalonada.

32

$$\begin{aligned} x_1 - 3x_2 + x_4 - 5x_5 &= -5 \\ -5x_2 + 4x_3 + x_5 &= -3 \\ -x_3 + x_4 &= -7 \end{aligned}$$

Solución *(a)* La matriz ampliada es en este caso

$$\left[\begin{array}{rrrrr|r} 1 & -3 & 0 & 1 & -5 & -5 \\ 0 & -5 & 4 & 0 & 1 & -3 \\ 0 & 0 & -1 & 1 & 0 & -7 \end{array}\right].$$

(b) Puesto que la matriz aumentada está en forma escalonada, el sistema está escalonado. ✔

33

$$\begin{aligned} 2x_1 - x_2 + 3x_4 - 5x_5 &= 2 \\ x_2 + 4x_3 - x_5 &= -4 \\ -2x_1 + 3x_4 &= -3 \end{aligned}$$

Solución *(a)* Para este sistema la matriz ampliada correspondiente es

$$\left[\begin{array}{rrrrr|r} 2 & -1 & 0 & 3 & -5 & 2 \\ 0 & 1 & 4 & 0 & -1 & -4 \\ -2 & 0 & 0 & 3 & 0 & -3 \end{array}\right].$$

(b) Dado que en este sistema la matriz aumentada no está en forma escalonada, el sistema no está escalonado. ✔

En los ejercicios 34 a 36, determinar si el sistema escalonado dado tiene solución única, tiene una infinidad de soluciones o es inconsistente, sin resolver dicho sistema; y en los sistemas que sean consistentes indicar las variables libres y las variables ligadas.

34

$$\left[\begin{array}{rrrrr|r} 3 & -1 & -1 & -3 & 5 & 1 \\ 0 & 2 & -4 & 0 & 2 & -2 \\ 0 & 0 & 0 & 3 & 0 & -3 \end{array}\right]$$

Solución El sistema es consistente porque no tiene una fila nula a la izquierda de la partición con un correspondiente registro distinto de cero a la derecha. Variables libres: x_3 y x_5 (que corresponden a las columnas sin pivote). Variables ligadas x_1, x_2 y x_4 (que corresponden a columnas con pivote). Como el sistema es consistente y tiene variables libres (equivalentemente columnas con pivote), tiene entonces una infinidad de soluciones. ✔

35

$$\left[\begin{array}{rrrrr|r} 1 & -1 & 2 & 4 & -5 & 8 \\ 0 & 1 & -2 & 0 & 3 & -5 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{array}\right]$$

Solución El sistema es inconsistente porque tiene una fila nula a la izquierda de la partición con un registro correspondiente distinto de cero a la derecha. ✔

36

$$\left[\begin{array}{rrrrr|r} 1 & -1 & 0 & 2 & 2 & 8 \\ 0 & -3 & 2 & 4 & 6 & 9 \\ 0 & 0 & 2 & -3 & 5 & 9 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Solución El sistema es consistente porque la fila nula a la izquierda de la partición tiene registro también nulo a la derecha de la partición. Variables ligadas: x_1, x_2, x_3, x_4 y x_5 (las que corresponden a columnas con pivote). No tiene variables libres. Puesto que el sistema es consistente y no tiene variables libres tiene solución única. ✔

37 Resolver el sistema escalonado del ejercicio 34 haciendo sustitución regresiva.

Solución La última fila de la matriz ampliada equivale a $3x_4 = -3$; por tanto, $x_4 = -1$. La segunda fila de la matriz ampliada equivale a $2x_2 - 4x_3 + 2x_5 = -2$ (con x_3 y x_5 variables libres), que al despejar la variable ligada x_2 produce $x_2 = -1 - x_5 + 2x_3$. La primera fila de la matriz ampliada equivale a $3x_1 - x_2 - x_3 - 3x_4 + 5x_5 = 1$, con x_3 y x_5 variables libres; que al despejar la variable ligada x_1 produce $x_1 = (1 + x_2 + x_3 + 3x_4 - 5x_5)/3 = (1 - 1 - x_5 + 2x_3 + x_3 - 3 - 5x_5)/3 = x_3 - 2x_5 - 1$. Entonces la solución está dada por

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} s - 2r - 1 \\ -1 - r + 2s \\ s \\ -1 \\ r \end{bmatrix}; \quad r, s \in \mathbb{R}.$$ ✔

38 Resolver el sistema escalonado del ejercicio 36 haciendo sustitución regresiva.

Solución La quinta fila de la ampliación equivale a $x_5 = 1$; la cuarta fila de la ampliación equivale a $x_4 = 2$; la tercera fila de la ampliación equivale a $2x_3 - 3x_4 + 6x_5 = 9$, que al despejar la variable ligada x_3 produce $x_3 = (9 + 3x_4 - 5x_5)/2 = (9 + 6 - 5)/2 = 5$; la segunda fila de la matriz ampliada equivale a $-3x_2 + 2x_3 + 4x_4 + 6x_5 = 9$, que al despejar la variable ligada x_2 produce $x_2 = (9 - 2x_3 - 4x_4 - 6x_5)/(-3) = (9 - 10 - 8 - 6)/(-3) = 5$; la primera fila de la matriz ampliada equivale a $x_1 - x_2 + 2x_4 + 2x_5 = 8$, que al despejar la variable ligada x_1 produce $x_1 = 8 + x_2 - 2x_4 - 2x_5 = 8 + 5 - 4 - 2 = 7$.

Así la solución es $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \\ 5 \\ 2 \\ 1 \end{bmatrix}$. ✔

39 Encontrar una matriz H que esté en forma escalonada y que sea equivalente por filas a la matriz

$$A = \begin{bmatrix} -3 & 2 & 1 & 4 & 4 \\ 2 & -1 & 5 & 6 & -3 \\ 7 & 2 & 1 & -1 & 5 \\ 1 & -1 & 2 & 3 & 4 \end{bmatrix}$$ por medio del método de Gauss.

Solución

$$\begin{bmatrix} -3 & 2 & 1 & 4 & 4 \\ 2 & -1 & 5 & 6 & -3 \\ 7 & 2 & 1 & -1 & 5 \\ 1 & -1 & 2 & 3 & 4 \end{bmatrix} \xleftrightarrow{R_1 \leftrightarrow R_4} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 2 & -1 & 5 & 6 & -3 \\ 7 & 2 & 1 & -1 & 5 \\ -3 & 2 & 1 & 4 & 4 \end{bmatrix}$$

$$\xleftrightarrow{\substack{R_2 \leftrightarrow -2R_1+R_2 \\ R_3 \leftrightarrow -7R_1+R_3 \\ R_4 \leftrightarrow \ \ 3R_1+R_4}} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 0 & -11 \\ 0 & 9 & -13 & -22 & -23 \\ 0 & -1 & 7 & 13 & 16 \end{bmatrix}$$

$$\xleftrightarrow{\substack{R_3 \leftrightarrow -9R_2+R_3 \\ R_4 \leftrightarrow \ \ R_2+R_4}} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 0 & -11 \\ 0 & 0 & -22 & -22 & 76 \\ 0 & 0 & 8 & 13 & 5 \end{bmatrix}$$

$$\xleftrightarrow{R_3 \leftrightarrow (1/22)R_3} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 0 & -11 \\ 0 & 0 & -1 & -1 & 38/11 \\ 0 & 0 & 8 & 13 & 5 \end{bmatrix}$$

$$\xleftrightarrow{R_4 \leftrightarrow 8R_3+R_4} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 0 & -11 \\ 0 & 0 & -1 & -1 & 38/11 \\ 0 & 0 & 0 & 5 & 359/11 \end{bmatrix} = H$$

La matriz H está en forma escalonada y es equivalente a la matriz A. ✔

40 Resolver por el método de Gauss el sistema

$$\begin{aligned} x_1 - x_2 + 3x_3 + 5x_4 - 6x_5 &= -15 \\ 3x_1 - 4x_2 + 5x_3 - 2x_4 + x_5 &= -43 \\ -2x_1 + 3x_2 - x_3 + 3x_4 - 5x_5 &= 17 \\ -x_1 - x_2 + 2x_3 - 2x_4 + 3x_5 &= -13 \\ 5x_1 - 3x_2 + x_3 - 4x_4 + 7x_5 &= -20 \\ 4x_1 - 5x_2 + 8x_3 + 3x_4 - 5x_5 &= -58 \end{aligned}$$

Solución Primero se lleva la matriz ampliada del sistema a forma escalonada.

$$\left[\begin{array}{rrrrr|r} 1 & -1 & 3 & 5 & -6 & -15 \\ 3 & -4 & 5 & -2 & 1 & -43 \\ -2 & 3 & -1 & 3 & -5 & 17 \\ -1 & -1 & 2 & -2 & 3 & -13 \\ 5 & -3 & 1 & -4 & 7 & -20 \\ 4 & -5 & 8 & 3 & -5 & -58 \end{array}\right] \sim \left[\begin{array}{rrrrr|r} 1 & -1 & 3 & 5 & -6 & -15 \\ 0 & -1 & -4 & -17 & 19 & 2 \\ 0 & 1 & 5 & 13 & -17 & -13 \\ 0 & -2 & 5 & 3 & -3 & -28 \\ 0 & 2 & -14 & -29 & 37 & 55 \\ 0 & -1 & -4 & -17 & 19 & 2 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrrr|r} 1 & -1 & 3 & 5 & -6 & -15 \\ 0 & -1 & -4 & -17 & 19 & 2 \\ 0 & 0 & 1 & -4 & 2 & -11 \\ 0 & 0 & 13 & 37 & -41 & -32 \\ 0 & 0 & -22 & -63 & 75 & 59 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrrr|r} 1 & -1 & 3 & 5 & -6 & -15 \\ 0 & -1 & -4 & -17 & 19 & 2 \\ 0 & 0 & 1 & -4 & 2 & -11 \\ 0 & 0 & 0 & 89 & -67 & 111 \\ 0 & 0 & 0 & -151 & 119 & -183 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrrr|r} 1 & -1 & 3 & 5 & -6 & -15 \\ 0 & -1 & -4 & -17 & 19 & 2 \\ 0 & 0 & 1 & -4 & 2 & -11 \\ 0 & 0 & 0 & 89 & -67 & 111 \\ 0 & 0 & 0 & 0 & 474 & 474 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Después se hace sustitución regresiva:

$$x_5 = 1; x_4 = (111 + 67x_5)/89 = 2; \quad x_3 = -11 + 4x_4 - 2x_5 = -5;$$

$$x_2 = -2 - 4x_3 - 17x_4 + 19x_5 = 3; \quad x_1 = -15 + x_2 - 3x_3 - 5x_4 + 6x_5 = -1$$

El sistema tiene solución única: $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -5 \\ 2 \\ 1 \end{bmatrix}$. ✔

41 Resolver el sistema

$$\begin{aligned} x_1 + x_2 + 2x_3 - x_4 &= 10 \\ x_1 - x_2 + x_3 - x_4 &= 9 \\ x_1 + x_2 - x_3 - x_4 &= 1 \\ -x_1 - 2x_2 + x_3 + x_4 &= 2 \end{aligned}$$

por el método de Gauss.

Solución Se lleva la matriz aumentada del sistema a forma escalonada

$$\left[\begin{array}{rrrr|r} 1 & 1 & 2 & -1 & 10 \\ 1 & -1 & 1 & -1 & 9 \\ 1 & 1 & -1 & -1 & 1 \\ -1 & -2 & 1 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{rrrr|r} 1 & 1 & 2 & -1 & 10 \\ 0 & -2 & -1 & 0 & -1 \\ 0 & 0 & -3 & 0 & -9 \\ 0 & -1 & 3 & 0 & 12 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 1 & 2 & -1 & 10 \\ 0 & -2 & -1 & 0 & -1 \\ 0 & 0 & -3 & 0 & -9 \\ 0 & 0 & -7 & 0 & -25 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 1 & 2 & -1 & 10 \\ 0 & -2 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & -7 & 0 & -25 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 1 & 2 & -1 & 10 \\ 0 & -2 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & -4 \end{array}\right]$$

y se observa que el sistema es inconsistente porque, en la forma escalonada equivalente, la última fila del lado izquierdo de la partición es nula y el registro correspondiente del lado derecho es distinto de cero. ✔

42 Resolver el sistema

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= -2 \\ -3x_1 + 5x_2 - x_3 + 2x_4 &= 6 \\ 4x_1 + 3x_2 + -5x_3 - x_4 &= -11 \\ -7x_1 + 19x_2 - 9x_3 + 8x_4 &= 14 \end{aligned}$$

mediante el método de Gauss.

Solución Primero se obtiene una forma escalonada de la matriz ampliada del sistema.

$$\left[\begin{array}{rrrr|r} 1 & 2 & -3 & 1 & -2 \\ -3 & 5 & -1 & 2 & 6 \\ 4 & 3 & -5 & -1 & -11 \\ -7 & 19 & -9 & 8 & 14 \end{array}\right] \sim \left[\begin{array}{rrrr|r} 1 & 2 & -3 & 1 & -2 \\ 0 & 11 & -10 & 5 & 0 \\ 0 & -5 & 7 & -5 & -3 \\ 0 & 33 & -30 & 15 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 2 & -3 & 1 & -2 \\ 0 & 1 & 4 & -5 & -6 \\ 0 & -5 & 7 & -5 & -3 \\ 0 & 11 & -10 & 5 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 2 & -3 & 1 & -2 \\ 0 & 1 & 4 & -5 & -6 \\ 0 & 0 & 27 & -30 & -33 \\ 0 & 0 & -54 & 60 & 66 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 2 & -3 & 1 & -2 \\ 0 & 1 & 4 & -5 & -6 \\ 0 & 0 & 27 & -30 & -33 \\ 0 & 0 & -54 & 60 & 66 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 2 & -3 & 1 & -2 \\ 0 & 1 & 4 & -5 & -6 \\ 0 & 0 & 27 & -30 & -33 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 2 & -3 & 1 & -2 \\ 0 & 1 & 4 & -5 & -6 \\ 0 & 0 & 9 & -10 & -11 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Después se hace sustitución regresiva:

$x_3 = (-11/9) + (10/9)x_4$; $x_2 = -6 - 4x_3 + 5x_4 = (-10/9) + (5/9)x_4$; $x_1 = (-31/9) + (11/9)x_4$;

es decir,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -\frac{31}{9} + \frac{11}{9}r \\ -\frac{10}{9} + \frac{5}{9}r \\ -\frac{11}{9} + \frac{10}{9}r \\ r \end{bmatrix}, \quad r \in \mathbb{R}. \checkmark$$

43 Resolver los siguientes sistemas que tienen la misma matriz de coeficientes.

(a)
$$\begin{aligned} 3x_1 - x_2 + 5x_3 + 2x_4 + x_5 &= 17 \\ -x_1 + x_2 - 2x_3 - x_4 + 3x_5 &= -2 \\ 2x_1 + 2x_2 + 4x_3 + x_4 - 4x_5 &= -4 \end{aligned}$$

(b)
$$\begin{aligned} 3y_1 - y_2 + 5y_3 + 2y_4 + y_5 &= 7 \\ -y_1 + y_2 - 2y_3 - y_4 + y_5 &= -6 \\ 2y_1 + 2y_2 + 4y_3 + y_4 - 4y_5 &= 10 \end{aligned}$$

(c)
$$\begin{aligned} 3z_1 - z_2 + 5z_3 + 2z_4 + z_5 &= 6 \\ -z_1 + z_2 - 2z_3 - z_4 + 3z_5 &= -5 \\ 2z_1 + 2z_2 + 4z_3 + z_4 - 4z_5 &= 12 \end{aligned}$$

Solución Se lleva a forma escalonada la matriz ampliada del sistema.

$$\left[\begin{array}{rrrrr|rrr} 3 & -1 & 5 & 2 & 1 & 17 & 7 & 6 \\ -1 & 1 & -2 & -1 & 3 & -2 & -6 & -5 \\ 2 & 2 & 4 & 1 & -4 & -4 & 10 & 12 \end{array}\right] \sim$$

$$\left[\begin{array}{rrrrr|rrr} -1 & 1 & -2 & -1 & 3 & -2 & -6 & -5 \\ 3 & -1 & 5 & 2 & 1 & 17 & 7 & 6 \\ 2 & 2 & 4 & 1 & -4 & -4 & 10 & 12 \end{array}\right] \sim$$

$$\left[\begin{array}{rrrrr|rrr} -1 & 1 & -2 & -1 & 3 & -2 & -6 & -5 \\ 0 & 2 & -1 & -1 & 10 & 11 & -11 & -9 \\ 0 & 4 & 0 & -1 & 2 & -8 & -2 & 2 \end{array}\right] \sim$$

$$\left[\begin{array}{rrrrr|rrr} -1 & 1 & -2 & -1 & 3 & -2 & -6 & -5 \\ 0 & 2 & -1 & -1 & 10 & 11 & -11 & -9 \\ 0 & 0 & 2 & 1 & -18 & -30 & 20 & 20 \end{array}\right].$$

Se hace sustitución regresiva de cada término independiente.

(a) $x_3 = (-30 - x_4 + 18x_5)/2 = -15 - (1/2)x_4 + 9x_5$; $x_2 = (11 + x_3 + x_4 - 10x_5)/2 = -2 + (1/4)x_4 - (1/2)x_5$; $x_1 = 2 + x_2 - 2x_3 - x_4 + 3x_5 = (1/4)x_4 - (31/2)x_5$. Con lo que

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 30 + \frac{1}{4}s - \frac{31}{2}s \\ -2 + \frac{1}{4}s - \frac{1}{2}r \\ -15 - \frac{1}{2}s + 9r \\ s \\ r \end{bmatrix}; \quad r, s \in \mathbb{R}.$$

(b) $y_3 = (20 - y_4 + 18y_5)/2 = 10 - \frac{1}{2}s + 9r$; $y_2 = (-11 + y_3 + y_4 - 10y_5)/2 = -\frac{1}{2} + \frac{1}{4}s - \frac{1}{2}r$; $y_1 = 6 + y_2 - 2y_3 - y_4 + 3y_5 = -\frac{29}{2} + \frac{1}{4}s - \frac{31}{2}r$. Por lo que

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} -\frac{29}{2} + \frac{1}{4}s - \frac{31}{2}r \\ -\frac{1}{2} + \frac{1}{4}s - \frac{1}{2}r \\ 10 - \frac{1}{2}s + 9r \\ s \\ r \end{bmatrix}; \quad r, s \in \mathbb{R}.$$

(c) $z_3 = (20 - z_4 + 18z_5)/2 = 10 - \frac{1}{2}s + 9r$; $z_2 = (-9 + z_3 + z_4 - 10z_5)/2 = \frac{1}{2} + \frac{1}{4}s - \frac{1}{2}r$; $z_1 = 5 + z_2 - 2z_3 - z_4 + 3z_5 = -\frac{29}{2} + \frac{1}{4}s - \frac{31}{2}r$. Por tanto

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} -\frac{29}{2} + \frac{1}{4}s - \frac{31}{2}r \\ \frac{1}{2} + \frac{1}{4}s - \frac{1}{2}r \\ 10 - \frac{1}{2}s + 9r \\ s \\ r \end{bmatrix}; \quad r, s \in \mathbb{R}.$$ ✔

44 Encontrar la matriz en forma escalonada reducida equivalente a la matriz

$$A = \begin{bmatrix} 1 & -1 & 2 & 3 & 1 \\ 2 & -1 & 3 & 4 & -1 \\ -1 & 1 & 2 & 5 & 3 \\ 3 & 2 & -1 & 1 & 2 \end{bmatrix}$$

por el método de Gauss-Jordan.

Solución

$$\begin{bmatrix} 1 & -1 & 2 & 3 & 1 \\ 2 & -1 & 3 & 4 & -1 \\ -1 & 1 & 2 & 5 & 3 \\ 3 & 2 & -1 & 1 & 2 \end{bmatrix} \xleftrightarrow[\substack{R_3 \leftrightarrow R_1+R_3 \\ R_2 \leftrightarrow -3R_1+R_3}]{R_2 \leftrightarrow -2R_1+R_2} \begin{bmatrix} 1 & -1 & 2 & 3 & 1 \\ 0 & 1 & -1 & -2 & -3 \\ 0 & 0 & 4 & 8 & 4 \\ 0 & 5 & -7 & -8 & -1 \end{bmatrix}$$

$$\xleftrightarrow{R_4 \leftrightarrow -5R_2+R_4} \begin{bmatrix} 1 & -1 & 2 & 3 & 1 \\ 0 & 1 & -1 & -2 & -3 \\ 0 & 0 & 4 & 8 & 4 \\ 0 & 0 & -2 & 2 & 14 \end{bmatrix}$$

$$\xleftrightarrow{R_3 \leftrightarrow (1/4)R_3} \begin{bmatrix} 1 & -1 & 2 & 3 & 1 \\ 0 & 1 & -1 & -2 & -3 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & -2 & 2 & 14 \end{bmatrix}$$

$$\xleftrightarrow{R_4 \leftrightarrow 2R_3+R_4} \begin{bmatrix} 1 & -1 & 2 & 3 & 1 \\ 0 & 1 & -1 & -2 & -3 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 6 & 16 \end{bmatrix}$$

$$\xleftrightarrow{R_4 \leftrightarrow (1/6)R_4} \begin{bmatrix} 1 & -1 & 2 & 3 & 1 \\ 0 & 1 & -1 & -2 & -3 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & \frac{8}{3} \end{bmatrix}$$

$$\xleftrightarrow[\substack{R_2 \leftrightarrow 2R_4+R_2 \\ R_1 \leftrightarrow -3R_4+R_1}]{R_3 \leftrightarrow -2R_4+R_3} \begin{bmatrix} 1 & -1 & 2 & 0 & -7 \\ 0 & 1 & -1 & 0 & \frac{7}{3} \\ 0 & 0 & 1 & 0 & -\frac{13}{3} \\ 0 & 0 & 0 & 1 & \frac{8}{3} \end{bmatrix}$$

$$\xleftrightarrow[R_1 \leftrightarrow -2R_3+R_1]{R_2 \leftrightarrow R_3+R_2} \begin{bmatrix} 1 & -1 & 0 & 0 & \frac{5}{3} \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & -\frac{13}{3} \\ 0 & 0 & 0 & 1 & \frac{8}{3} \end{bmatrix}$$

$$\xleftrightarrow{R_1 \leftrightarrow R_2+R_1} \begin{bmatrix} 1 & 0 & 0 & 0 & -\frac{1}{3} \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & -\frac{13}{3} \\ 0 & 0 & 0 & 1 & \frac{8}{3} \end{bmatrix} = H.$$

La matriz H está en forma escalonada reducida y $A \sim H$. ✔

45 Resolver el siguiente sistema por el método de Gauss-Jordan.

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 0 \\ 2x_1 - x_2 + x_3 &= -2 \\ -x_1 + 3x_2 - 2x_3 &= 2 \\ x_1 + x_2 + 2x_3 &= 2 \end{aligned}$$

Solución Se lleva primero la matriz ampliada del sistema a forma escalonada reducida mediante el método de Gauss-Jordan:

$$\left[\begin{array}{rrr|r} 1 & -1 & 2 & 0 \\ 2 & -1 & 1 & -2 \\ -1 & 3 & -2 & 2 \\ 1 & 1 & 2 & 2 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & -1 & 2 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 2 & 0 & 2 \\ 0 & 2 & 0 & 2 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & -1 & 2 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 6 & 6 \\ 0 & 0 & 6 & 6 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|r} 1 & -1 & 2 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 6 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & -1 & 2 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|r} 1 & -1 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Después se hace sustitución regresiva.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}. \checkmark$$

46 Resolver el siguiente sistema homogéneo por el método de Gauss.

$$\begin{aligned} 7x_1 - x_2 + x_3 - x_4 &= 0 \\ 15x_1 + 3x_2 - x_3 + x_4 &= 0 \\ 4x_1 - 2x_2 + 7x_3 - 5x_4 &= 0 \end{aligned}$$

Solución Se lleva la matriz de coeficientes del sistema a forma escalonada utilizando el método de Gauss:

$$\begin{bmatrix} 7 & -1 & 1 & -1 \\ 15 & 3 & -1 & 1 \\ 4 & -2 & 7 & -5 \end{bmatrix} \sim \begin{bmatrix} 7 & -1 & 1 & -1 \\ 1 & 5 & -3 & 3 \\ 4 & -2 & 7 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & 3 \\ 7 & -1 & 1 & -1 \\ 4 & -2 & 7 & -5 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 5 & -3 & 3 \\ 0 & -36 & 22 & -22 \\ 0 & -22 & 19 & -17 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & 3 \\ 0 & 1 & -\frac{11}{18} & \frac{11}{18} \\ 0 & -22 & 19 & -17 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 5 & -3 & 3 \\ 0 & 1 & -\frac{11}{18} & \frac{11}{18} \\ 0 & -22 & 19 & -17 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & 3 \\ 0 & 1 & -\frac{11}{18} & \frac{11}{18} \\ 0 & 0 & \frac{50}{9} & -\frac{32}{9} \end{bmatrix}$$

Se hace sustitución regresiva (igualando las ecuaciones a 0 porque el sistema es homogéneo):

$$x_3 = \frac{16}{25}x_4; \quad x_2 = \frac{11}{18}x_3 - \frac{11}{18}x_4 = -\frac{11}{50}x_4; \quad x_1 = -5x_2 + 3x_3 - 3x_4 = \frac{1}{50}x_4.$$

Por tanto,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{50}r \\ -\frac{11}{50}r \\ \frac{16}{25}r \\ r \end{bmatrix}; \quad r \in \mathbb{R}. \text{ O, equivalentemente,}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} s \\ -11s \\ 32s \\ 50s \end{bmatrix}; \quad s \in \mathbb{R}. \checkmark$$

47 Resolver el siguiente sistema y de manera simultánea encontrar la solución del sistema homogéneo asociado.

$$\begin{aligned} x_1 - 2x_2 + 3x_3 - 4x_4 &= -1 \\ -x_1 + 3x_2 - 5x_3 + x_4 &= 2 \\ 2x_1 - x_2 + 2x_3 - 3x_4 &= 3 \\ 4x_1 - 5x_2 + 8x_3 - 11x_4 &= 1 \end{aligned}$$

Solución Se resuelve el sistema no homogéneo por el método de Gauss llevando la matriz ampliada a forma escalonada.

$$\left[\begin{array}{rrrr|r} 1 & -2 & 3 & -4 & -1 \\ -1 & 3 & -5 & 1 & 2 \\ 2 & -1 & 2 & -3 & 3 \\ 4 & -5 & 8 & -11 & 1 \end{array}\right] \sim \left[\begin{array}{rrrr|r} 1 & -2 & 3 & -4 & -1 \\ 0 & 1 & -2 & -3 & 1 \\ 0 & 3 & -4 & 5 & 5 \\ 0 & 3 & -4 & 5 & 5 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & -2 & 3 & -4 & -1 \\ 0 & 1 & -2 & -3 & 1 \\ 0 & 3 & -4 & 5 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|r} 1 & -2 & 3 & -4 & -1 \\ 0 & 1 & -2 & -3 & 1 \\ 0 & 0 & 2 & 14 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Se hace sustitución regresiva: $x_3 = 1 - 7x_4$; $x_2 = 1 + 2x_3 + 3x_4 = 3 - 11x_4$; $x_1 = -1 + 2x_2 - 3x_3 + 4x_4 = 2 + 3x_4$. Por tanto,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2+3r \\ 3-11r \\ 1-7r \\ r \end{bmatrix}; \quad r \in \mathbb{R},$$

es la solución para el sistema $A\vec{x} = \vec{0}$. Para encontrar la solución del sistema homogéneo asociado $A\vec{x} = \vec{0}$ se escribe la solución precedente como (cfr. 1.2.7, pág. 32)

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 3r \\ -11r \\ -7r \\ r \end{bmatrix}; \quad r \in \mathbb{R},$$

y se obtiene una solución particular del sistema no homogéneo y la solución del sistema homogéneo asociado

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3r \\ -11r \\ -7r \\ r \end{bmatrix}; \quad r \in \mathbb{R}.$$ ✔

48 Determinar los valores de α para que el sistema

$$\begin{aligned} x_1 - x_2 + \alpha x_3 &= -2 \\ -x_1 + 2x_2 - \alpha x_3 &= 3 \\ \alpha x_1 + x_2 + x_3 &= 2 \end{aligned}$$

tenga solución única.

Solución Al llevar la matriz ampliada del sistema a forma escalonada se obtiene

$$\left[\begin{array}{rrr|r} 1 & -1 & \alpha & -2 \\ -1 & 2 & -\alpha & 3 \\ \alpha & 1 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{lll|c} 1 & -1 & \alpha & -2 \\ 0 & 1 & 0 & 1 \\ 0 & \alpha+1 & 1-\alpha^2 & 2+2\alpha \end{array}\right]$$
$$\sim \left[\begin{array}{lll|c} 1 & -1 & \alpha & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1-\alpha^2 & 1+\alpha \end{array}\right]$$

así que el sistema tiene solución única para toda $\alpha \neq \pm 1$. (Si $\alpha = 1$, el sistema es inconsistente; y si $\alpha = -1$, el sistema tiene una infinidad de soluciones.) ✔

49 Determine los valores de α para que el sistema homogéneo

$$\begin{aligned} \alpha x - y + z &= 0 \\ x + 2y - \alpha z &= 0 \\ x + 2y - z &= 0 \end{aligned}$$

tenga soluciones distintas de cero.

Solución
$$\begin{bmatrix} \alpha & -1 & 1 \\ 1 & 2 & -\alpha \\ 1 & 2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -\alpha \\ \alpha & -1 & 1 \\ 1 & 2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -\alpha \\ 0 & -2\alpha-1 & \alpha^2+1 \\ 0 & 0 & \alpha-1 \end{bmatrix}.$$

Dado que el sistema homogéneo es cuadrado, tiene solución no trivial si la matriz de coeficientes no es equivalente a la identidad (cfr. teorema 1.6); y esto sucede si $\alpha = 1$ o $\alpha = -1/2$. ✔

50 Resolver el sistema

$$\begin{aligned} iz_1 - iz_2 + (3+i)z_3 &= 4+2i \\ (1+i)z_1 - z_2 + z_3 &= 2 \\ 3iz_1 + z_2 - 3z_3 &= -3+4i \end{aligned}$$

Solución
$$\left[\begin{array}{ccc|c} i & -i & 3+i & 4+2i \\ 1+i & -1 & 1 & 2 \\ 3i & 1 & -3 & -3+4i \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -1 & 1-3i & 2-4i \\ 1+i & -1 & 1 & 2 \\ 3i & 1 & -3 & -3+4i \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|c} 1 & -1 & 1-3i & 2-4i \\ 0 & i & -3+2i & -4+2i \\ 0 & 1+3i & -12-3i & -15-2i \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|c} 1 & -1 & 1-3i & 2-4i \\ 0 & 1 & 2+3i & 2+4i \\ 0 & 1+3i & -12-3i & -15-2i \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|c} 1 & -1 & 1-3i & 2-4i \\ 0 & 1 & 2+3i & 2+4i \\ 0 & 0 & -5-12i & -5-12i \end{array}\right]$$

de donde, al hacer sustitución regresiva,

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 1 \\ i \\ 1 \end{bmatrix}$$

es la solución del sistema. ✔

1.3.2 Ejercicios propuestos

El lector encontrará la respuesta a los ejercicios en cursiva en el apéndice E al final del libro.

Matrices (respuestas en páginas 1073-1074)

En los ejercicios 1 a 3 sean

$$A=\begin{bmatrix}-1 & 2 & -3\\ 2 & 5 & -4\end{bmatrix},\quad B=\begin{bmatrix}5 & 0 & -1\\ -2 & 3 & -4\end{bmatrix},\quad C=\begin{bmatrix}-1 & 0 & -2 & 4\\ 3 & 7 & -3 & 2\\ 4 & -6 & -3 & 2\end{bmatrix}\quad \text{y}\quad D=\begin{bmatrix}2\\ 3\\ -6\end{bmatrix}.$$

1 Calcular *(a)* $A+B$, *(b)* $A+C$, *(c)* $3A-5B$, *(d)* $6D$.

2 Hallar *(a)* AB, *(b)* AC, *(c)* AD, *(d)* BC, *(e)* BD, *(f)* CD.

3 Encontrar *(a)* A^t, *(b)* A^tC, *(c)* D^tA^t, *(d)* B^tA, *(e)* D^tD, *(f)* DD^t.

En los ejercicios 4 a 7 sean

$$A=\begin{bmatrix}-8 & -5 & -7\\ -3 & 9 & 5\end{bmatrix},\quad B=\begin{bmatrix}7 & 5 & 4\\ 3 & 5 & -5\end{bmatrix},\quad C=\begin{bmatrix}4 & -8\\ -3 & 2\end{bmatrix},$$

$$D=\begin{bmatrix}3 & -6\\ 7 & 6\end{bmatrix}\quad \text{y}\quad E=\begin{bmatrix}-2\\ 4\end{bmatrix}.$$

4 Hallar *(a)* $-2A$, *(b)* $A+B$, *(c)* $B-2A$.

5 Calcular *(a)* $3C-E$, *(b)* AC, *(c)* CB.

6 Encontrar *(a)* CD, *(b)* EB, *(c)* E^tB, *(d)* AA^t.

7 Calcular *(a)* A^tB, *(b)* B^tA, *(c)* BB^t, *(e)* A^tA.

En los ejercicios 8 a 15 sean

$$A=\begin{bmatrix}5 & -5\\ 9 & -1\\ 0 & -12\end{bmatrix},\quad B=\begin{bmatrix}-8 & 3\\ -6 & -2\end{bmatrix},\quad C=\begin{bmatrix}-1 & 4 & 8\\ -6 & 2 & -4\end{bmatrix},$$

$$D=\begin{bmatrix}1 & -4 & -1\\ -4 & 1 & 4\\ -8 & -9 & 5\end{bmatrix}\quad \text{y}\quad E=\begin{bmatrix}9 & 0 & 8\\ -3 & 8 & 4\\ 7 & 1 & 7\end{bmatrix}.$$

8 Calcular *(a)* $D+E$, *(b)* $D-E$, *(c)* $4A$, *(d)* $-8C$.

9 Encontrar *(a)* $B-2C$, *(b)* $3E-5D$, *(c)* $-4(D+3E)$, *(d)* $C-C$.

10 Hallar *(a)* $-A^t+3C$, *(b)* E^t-D^t, *(c)* $(D-E)^t$, *(d)* B^t+4C^t.

11 Calcular *(a)* $\frac{1}{4}C^t-\frac{1}{2}A$, *(b)* B^t-B, *(c)* $4E^t-3D^t$, *(d)* $(3E^t-2D^t)^t$.

12 Calcular *(a)* AB, *(b)* BA, *(c)* $(2E)D$, *(d)* $(AB)C$.

13 Hallar *(a)* $A(BC)$, *(b)* CC^t, *(c)* $(DA)^t$, *(d)* $(C^tB)A^t$.

14 Encontrar *(a)* $(3D^t-E)A$, *(b)* $(5B)C+3B$, *(c)* $(-AC)^t+5D^t$.

15 Calcular *(a)* $(BA^t-4C)^t$, *(b)* $B^t(CC^t-A^tA)$, *(c)* $D^tE^t-(ED)^t$.

16 Sea $A=\begin{bmatrix}0 & 0 & -3\\ 0 & 2 & 0\\ 4 & 0 & 0\end{bmatrix}$: *(a)* Hallar A^2; *(b)* Calcular A^8.

17 Sea $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$: *(a)* Calcular A^2; *(b)* Encontrar A^7.

18 Sean $A = \begin{bmatrix} -1 & 0 & 5 \end{bmatrix}$ y $\vec{x} = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}$. Calcular *(a)* $A\vec{x}$, *(b)* $\vec{x}A$.

19 Si $A = \begin{bmatrix} 0 & 1 \\ 0 & -2 \end{bmatrix}$, encontrar todas las matrices $B \in \mathfrak{M}_{2\times 2}$ tales que *(a)* $AB = \mathcal{O}$; *(b)* $BA = \mathcal{O}$; *(c)* $AB = \mathcal{O}$ y $BA = \mathcal{O}$.

20 Encontrar los valores de α, β, γ y δ tales que

$$\begin{bmatrix} \alpha & \beta & \gamma & \delta \\ 1 & 4 & 9 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 3 & 3 \\ 1 & 4 & 9 & 2 \end{bmatrix}.$$

21 Encontrar los valores de α y β tales que las matrices

$$A = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & \beta \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 3 & 0 & 3 \\ 0 & 1 & 0 \\ 3 & 0 & 3 \end{bmatrix}$$

conmuten; es decir, $AB = BA$.

22 Sea $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$: *(a)* Calcular A^2. *(b)* Encontrar A^3. *(c)* Determinar A^n para todo $n \in \mathbb{N}$.

23 Si $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$, hallar A^2 y calcular A^n para todo $n \in \mathbb{N}$.

24 Sea

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}:$$

(a) Calcular A^2. *(b)* Encontrar A^3. *(c)* Hallar A^4. *(d)* Conjeturar una fórmula para A^n, $n \in \mathbb{N}$, y demostrarla por inducción.

25 Encontrar todas las matrices $A \in \mathfrak{M}_{2\times 2}$ tales $A^2 = \mathcal{O}$.

26 Probar que una matriz A de tamaño 2×2 conmuta con toda matriz B del mismo tamaño ($AB = BA$) si y sólo si A conmuta con las matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Encontrar todas esas matrices A.

27 Encontrar todas las matrices A de tamaño 2×2 tales que $A^2 = I_2$, donde I_2 es la matriz identidad de orden 2.

28 Llenar las entradas vacías de la matriz 4×4

$$\begin{bmatrix} 2 & -3 & \square & 4 \\ \square & -2 & \square & 5 \\ -1 & 4 & -2 & \square \\ \square & \square & 5 & 7 \end{bmatrix}$$

de tal manera que se obtenga una matriz simétrica.

29 Llenar las entradas vacías de la matriz 4×4

$$\begin{bmatrix} 1 & -2 & \square & 3 \\ \square & 3 & \square & 6 \\ 1 & -2 & 4 & \square \\ \square & \square & 1 & 5 \end{bmatrix}$$

de tal manera que se obtenga una matriz simétrica.

30 Probar que si $A \in \mathfrak{M}_{m\times n}$, entonces $B = A + A^t$ es una matriz simétrica.

31 Demostrar que si A es una matriz cuadrada, entonces $(A^2)^t = (A^t)^2$ y $(A^3)^t = (A^t)^3$. (Indicación: Utilizar la propiedad 3*(b)* de la pág. 10.)

En los ejercicios 32 a 39 las afirmaciones dadas son falsas o verdaderas; si la afirmación es verdadera se debe demostrar con rigor su validez y si es falsa se debe dar un contraejemplo para mostrar que no es cierta. En cada caso se supone que las matrices involucradas A, B y C tienen los tamaños adecuados para efectuar las correspondientes operaciones.

32 $A = B \Rightarrow AC = BC$.

33 $A = B \Rightarrow CA = BC$.

34 $A = B \Rightarrow CA = CB$.

35 $A^2 = I_n \Rightarrow A = I_n$ o $A = -I_n$.

36 $AB = \mathscr{O} \Rightarrow A = \mathscr{O}$ o $B = \mathscr{O}$.

37 $C + A = B + A \Rightarrow C = B$.

38 $A^2 = I_n \Rightarrow A^m = I_n \; \forall n \geq 2$ (n entero).

39 Si A es una matriz cuadrada simétrica, $B = [b_{ij}]$ y $B = A^2$, entonces $b_{ii} \geq 0 \; \forall i$.

Sistemas lineales (respuestas en páginas 1072-1075)

40 Determinar cuáles de los siguientes sistemas de ecuaciones son lineales.

(a)
$$\begin{aligned} x_1 - x_2 + \sqrt{2}x_3 &= \pi \\ -x_1 + x_2 + ex_3 &= e \end{aligned}$$

(b)
$$\begin{aligned} \cos(x_1) + x_1x_2 - x_3 &= 5 \\ x_1 + x_2 - x_3 &= 6 \end{aligned}$$

(c)
$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 1 \\ 2x_1 + x_2 - x_3 &= 7 \\ -3x_1^2 + x_2 - x_3 &= 2 \end{aligned}$$

(d)
$$\begin{aligned} x_1 - 3x_2 - 4x_4 &= 2 \\ 2x_1 - 3x_2 - 2x_3 &= 3 \\ x_1 + x_2 - x_3 &= -5 \end{aligned}$$

41 Sea el sistema

$$\begin{aligned} x_1 - x_2 + 2x_3 + x_4 + x_5 + 3x_6 &= -1 \\ -3x_1 + 2x_2 + 4x_3 - x_4 + x_5 + 2x_6 &= 2 \\ -2x_1 + x_2 + x_3 + 2x_4 - x_5 + x_6 &= -1 \\ x_1 - x_2 + x_3 - x_4 + x_5 - x_6 &= 1 \end{aligned}$$

Determinar si *(a)* $(-1,-1,0,-1,0,0)$; *(b)* $(-10,-14,-2,-6,-4,3)$; *(c)* $(-13,-16,-4,-5,0,3)$; *(d)* $(1,1,2,1,1,-2)$ son soluciones.

42 Indicar las matrices que están en forma escalonada. Para las que sí lo están, señalar los pivotes de cada fila; y para las que no, mencionar las propiedades que no se cumplen para ese fin.

(a) $\begin{bmatrix} 2 & 0 & -2 & 1 & 2 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$

(b) $\begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & -1 & 3 \\ 0 & 3 & 0 & 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} -1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

De los ejercicios 43 a 46:

(a) Escribir la matriz aumentada correspondiente.

(b) Utilizar el inciso *(a)* para determinar si el sistema está en forma escalonada.

(c) Para los sistemas que estén en forma escalonada indicar cuáles son las variables libres y cuáles las ligadas.

43 $$\begin{aligned} x_1+2x_2-x_3+x_4+2x_5 &= -8 \\ x_1-x_3+x_5 &= -8 \\ x_3+2x_4 &= 4 \end{aligned}$$

44 $$\begin{aligned} 3x_1-x_2+x_3 &= 14 \\ x_2-x_3 &= 16 \\ x_3 &= -4 \end{aligned}$$

45 $$\begin{aligned} x_1-x_2+x_3-x_4-x_5 &= -6 \\ x_2-x_3+x_5 &= -3 \\ 2x_5 &= -8 \end{aligned}$$

46 $$\begin{aligned} x_1-4x_2+5x_3-x_4+x_5 &= -7 \\ x_4-5x_5 &= 12 \\ 3x_5 &= -4 \end{aligned}$$

En los ejercicios 47 a 50, determinar si el sistema escalonado tiene solución única, tiene una infinidad de soluciones o es inconsistente, sin resolverlo; y en los sistemas que son consistentes, indicar las variables libres y las variables ligadas.

47 $$\left[\begin{array}{cccccc|c} 0 & -1 & 2 & -1 & 0 & 3 & -3 \\ 0 & 0 & 1 & 0 & -1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

48 $$\left[\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 0 & -1 & 1 & 2 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

49 $$\left[\begin{array}{cccccc|c} 2 & -1 & -2 & 3 & 0 & -3 & 3 \\ 0 & 2 & 1 & 0 & -1 & 1 & 4 \\ 0 & 0 & -1 & 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

50 $$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 2 \\ 0 & 0 & -2 & 4 & -1 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

51 Resolver mediante sustitución regresiva. *(a)* el sistema escalonado del ejercicio 47; *(b)* el sistema escalonado del ejercicio 49; *(c)* el sistema escalonado del ejercicio 50.

En los ejercicios 52 a 59 utilizar el método de Gauss para encontrar una matriz H equivalente a la matriz dada que esté en forma escalonada.

52 $$\begin{bmatrix} 0 & 1 & -1 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 & 3 \\ 0 & 4 & -1 & 2 & 1 & 6 \end{bmatrix}$$

53 $$\begin{bmatrix} 3 & -2 & 1 & 1 & 4 \\ 2 & -1 & 1 & 2 & 3 \\ -2 & 1 & 1 & 0 & 1 \\ 5 & -2 & 3 & 1 & 1 \end{bmatrix}$$

54 $\begin{bmatrix} -1 & 2 & 1 & 1 & 3 \\ 3 & -6 & 4 & -2 & -5 \end{bmatrix}$

55 $\begin{bmatrix} 0 & 0 & 4 \\ 0 & 1 & -3 \\ 0 & 2 & 4 \end{bmatrix}$

56 $\begin{bmatrix} 2 & -3 & 2 & 4 & -5 & 2 & 1 \\ 4 & -2 & 1 & -1 & 2 & 1 & 0 \\ 3 & -1 & 2 & 1 & 0 & 1 & 0 \\ 1 & -1 & 2 & 3 & -1 & 1 & 2 \\ 2 & 1 & -1 & 0 & 1 & 1 & 1 \end{bmatrix}$

57 $\begin{bmatrix} -1 & 2 & 3 & -1 & 1 & 0 \\ 2 & 4 & 3 & -1 & 2 & 1 \\ 3 & -1 & 2 & 1 & 1 & 8 \\ 4 & 7 & -4 & 2 & -2 & 4 \end{bmatrix}$

58 $\begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & -2 & 3 & 4 \\ 1 & 2 & -4 & 6 \\ 2 & -1 & 1 & 2 \\ 2 & -3 & 1 & 5 \\ -3 & 2 & 1 & 1 \end{bmatrix}$

59 $\begin{bmatrix} 2 & -3 & 6 & 2 & 3 & 4 \\ 3 & -1 & 1 & 2 & 1 & 0 \\ 7 & 3 & -2 & 5 & 2 & 6 \\ -3 & 1 & 1 & 2 & 1 & 0 \\ 5 & -1 & 1 & 1 & 1 & 2 \\ -5 & 1 & 0 & 1 & -2 & 3 \end{bmatrix}$

En los ejercicios 60 a 71, resolver el sistema por medio del método de Gauss.

60
$$\begin{aligned} 3x_1 + x_2 + x_3 - x_4 + 2x_5 &= 1 \\ 2x_1 - x_2 + 2x_3 - x_4 + x_5 &= 2 \\ 3x_1 + x_2 - 2x_3 - x_4 - 2x_5 &= 1 \end{aligned}$$

61
$$\begin{aligned} x_1 + 5x_2 + 4x_3 - 13x_4 &= 5 \\ 3x_1 - x_2 + 2x_3 + 5x_4 &= -3 \\ 2x_1 + 2x_2 + 3x_3 - 4x_4 &= 1 \end{aligned}$$

62
$$\begin{aligned} 2x_1 + x_2 - 3x_3 &= -5 \\ 3x_1 - 2x_2 + 2x_3 &= 9 \\ 5x_1 - 3x_2 - x_3 &= 6 \end{aligned}$$

63
$$\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 + x_5 + 3x_6 &= 1 \\ 2x_1 - x_2 + 2x_3 + x_4 + 3x_5 - 4x_6 &= 3 \\ 3x_1 - 2x_2 + 4x_3 + 2x_4 + x_5 + x_6 &= 18 \\ -3x_1 + x_2 - x_3 + 2x_4 + x_5 + x_6 &= -1 \end{aligned}$$

64
$$\begin{aligned} 2x_1 + 3x_2 - 2x_3 &= -6 \\ x_1 - 2x_2 + 3x_3 &= 8 \\ 4x_1 - x_2 + 4x_3 &= 2 \end{aligned}$$

65
$$\begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 &= -1 \\ 2x_1 + 5x_2 - 8x_3 + 5x_4 &= -4 \\ 3x_1 + 4x_2 - 5x_3 + 2x_4 &= -1 \end{aligned}$$

66
$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= -1 \\ 3x_1 - 2x_2 - x_3 &= 4 \\ 2x_1 - 5x_2 + 3x_3 &= -1 \\ x_1 + 4x_2 + 6x_3 &= -5 \end{aligned}$$

67
$$\begin{aligned} x_1 - x_2 + x_3 + 2x_4 &= 8 \\ -x_1 + x_2 + x_3 + 3x_4 &= 4 \\ x_1 - x_2 + 3x_3 + 7x_4 &= 2 \\ -x_1 + x_2 + x_3 - 2x_4 &= -6 \end{aligned}$$

68
$$\begin{aligned} 2x_1 - x_2 + x_3 &= -1 \\ -x_1 + x_2 + 2x_3 &= 6 \\ -2x_1 + x_2 + x_3 &= 5 \\ 6x_1 - 2x_2 + 8x_3 &= 8 \\ 4x_1 - 3x_2 + 3x_3 &= -1 \end{aligned}$$

69
$$\begin{aligned} 3x_1 - 4x_2 + 3x_3 - x_4 &= 0 \\ 2x_1 - x_2 + 3x_3 - 5x_4 &= -11 \\ x_1 + x_2 + x_3 - x_4 &= 2 \\ -2x_1 + 3x_2 - x_3 + 2x_4 &= 9 \end{aligned}$$

70
$$\begin{aligned} x_1 - 2x_2 + x_3 + 3x_4 + 2x_5 &= 8 \\ 2x_1 - 3x_2 + 5x_3 + 7x_4 + 3x_5 &= 21 \\ -x_1 - 7x_3 - 5x_4 &= -18 \\ 4x_1 - 5x_2 + 13x_3 + 15x_4 + 5x_5 &= 47 \end{aligned}$$

71
$$\begin{aligned} 2x_1 - x_2 + 3x_3 - x_4 &= 16 \\ -x_1 + 2x_2 + x_3 + 2x_4 &= 6 \\ 4x_1 - 2x_2 + x_3 + x_4 &= 13 \\ 2x_1 + 3x_2 - 4x_3 - 2x_4 &= -25 \\ -x_1 + 9x_2 - x_3 - 4x_4 &= -23 \end{aligned}$$

Los ejercicios 72 a 75 contienen sistemas lineales que tienen la misma matriz de coeficientes pero distintos términos independientes, resolverlos por el método de Gauss.

72 $$\begin{aligned} x_1 - 2x_2 &= 5, -1, -7 \\ 2x_1 - 3x_2 &= 8, -1, -11 \end{aligned}$$

73 $$\begin{aligned} x_1 - 2x_2 + x_3 &= 3, -4, 1 \\ 3x_1 - x_2 + 2x_3 &= 7, -3, -4 \\ 2x_1 - x_2 + x_3 &= 4, -3, -3 \end{aligned}$$

74 $$\begin{aligned} x_1 + x_2 - 2x_3 + x_4 + 2x_5 &= -1, -6, -4 \\ -x_1 + x_2 - 2x_3 + x_4 - x_5 &= -2, -10, -1 \\ 2x_1 + 3x_2 - x_3 + 2x_4 + x_5 &= -2, -4, -3 \end{aligned}$$

75 $$\begin{aligned} x_1 - x_2 + x_3 + 4x_4 &= 6, 3, -6 \\ 2x_1 + x_2 + 5x_3 + 2x_4 &= 3, 0, -3 \\ x_1 + x_2 + 5x_3 &= 0, -3, 0 \end{aligned}$$

En los ejercicios 76 a 85, encontrar la forma escalonada reducida que es equivalente a cada matriz por medio del método de Gauss-Jordan.

76 $$\begin{bmatrix} 3 & 6 & -3 & 6 & 3 \\ 2 & 4 & 1 & -2 & 3 \\ 3 & 6 & 2 & -6 & 5 \end{bmatrix}$$

77 $$\begin{bmatrix} 2 & 3 & -2 & 5 & 1 \\ 9 & -1 & 2 & 0 & 4 \\ 4 & -5 & 6 & 5 & 7 \end{bmatrix}$$

78 $$\begin{bmatrix} 1 & 3 & -1 & 2 \\ 4 & -10 & 6 & 2 \\ 8 & 2 & 2 & 10 \\ 0 & -11 & 5 & -3 \end{bmatrix}$$

79 $$\begin{bmatrix} 0 & -2 & -6 & 4 \\ 0 & 4 & -1 & 3 \\ 0 & 0 & -6 & -3 \\ 0 & 5 & -3 & 4 \end{bmatrix}$$

80 $$\begin{bmatrix} 2 & 1 & 4 \\ 1 & 3 & 2 \\ 5 & 0 & 10 \end{bmatrix}$$

81 $$\begin{bmatrix} 2 & 4 & -2 \\ 4 & 8 & 3 \\ 1 & 1 & -2 \end{bmatrix}$$

82 $$\begin{bmatrix} 0 & 2 & -1 & 3 \\ -1 & 1 & 2 & 0 \\ -1 & 3 & 1 & 3 \\ 1 & 5 & 5 & 9 \end{bmatrix}$$

83 $$\begin{bmatrix} 0 & 0 & 2 & -4 \\ 0 & 0 & 1 & 2 \\ 1 & 3 & 2 & -4 \end{bmatrix}$$

84 $$\begin{bmatrix} -1 & 3 & 0 & 1 & 4 \\ 1 & -3 & 0 & 0 & -1 \\ 3 & -9 & 2 & 4 & -1 \\ 0 & 0 & 1 & 3 & -4 \end{bmatrix}$$

85 $$\begin{bmatrix} 0 & 0 & 1 & 2 & -1 & 4 \\ 0 & 0 & 0 & 1 & -1 & 3 \\ 2 & 4 & -2 & 1 & 3 & -5 \end{bmatrix}$$

En los ejercicios 86 a 91 resolver los sistemas por el método de Gauss-Jordan.

86 $$\begin{aligned} x_1 - 2x_3 + x_4 &= 4 \\ 2x_1 - x_2 + x_3 - 3x_4 &= -2 \\ 9x_1 - 3x_2 - x_3 - 7x_4 &= 3 \end{aligned}$$

87 $$\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= -7 \\ 3x_1 + 6x_2 - 8x_3 - 2x_4 &= -9 \end{aligned}$$

88 $$\begin{aligned} 2x_1 + 5x_2 - 4x_3 &= -2 \\ 2x_1 + 4x_2 - x_3 &= 1 \\ 4x_1 - 2x_2 + 5x_3 &= 9 \end{aligned}$$

89 $$\begin{aligned} -2x_3 - x_4 &= -1 \\ 2x_1 + 5x_2 - 6x_3 &= 3 \\ -x_1 - 2x_2 + x_3 - x_4 &= -2 \\ 4x_1 + 10x_2 - 9x_3 + x_4 &= 7 \end{aligned}$$

90 $$\begin{aligned} x_1 - 3x_2 + x_3 + 2x_4 &= -12 \\ x_1 - 2x_2 + 2x_3 + 4x_4 &= -14 \\ -x_1 - 6x_2 - 3x_3 - 4x_4 &= -9 \\ 2x_1 - 6x_2 - 3x_3 - 2x_4 &= -7 \end{aligned}$$

91 $$\begin{aligned} x_1 - x_2 + 2x_3 + x_4 - 2x_5 &= 3 \\ x_1 - x_2 - 2x_3 + x_4 - 2x_5 &= 7 \\ -x_1 + x_2 - 6x_3 - x_4 + 2x_5 &= 1 \\ 2x_1 - 2x_2 - 12x_3 + 2x_4 - 4x_5 &= 22 \end{aligned}$$

En los ejercicios 92 a 95 resolver los sistemas homogéneos.

92 $$\begin{aligned} x_1 - x_2 + x_3 - 4x_4 + 2x_5 &= 0 \\ 2x_1 - 3x_2 + x_3 - x_4 - x_5 &= 0 \\ -3x_1 - 2x_2 - x_3 + 3x_4 + x_5 &= 0 \\ -x_1 - 4x_2 + x_3 - 5x_4 + 5x_5 &= 0 \end{aligned}$$

93 $$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ -x_1 + x_2 - x_3 &= 0 \\ 3x_1 + 2x_2 - 2x_3 &= 0 \end{aligned}$$

94 $$\begin{aligned} 2x_1 + x_2 + 3x_3 &= 0 \\ 2x_1 + 4x_2 &= 0 \\ x_2 + x_3 &= 0 \end{aligned}$$

95 $$\begin{aligned} 4x_2 + 4x_3 + 2x_4 &= 0 \\ x_1 + 3x_2 + x_4 &= 0 \\ x_1 - 2x_2 - x_3 + x_4 &= 0 \end{aligned}$$

En los ejercicios 96 a 101, resolver el sistema $A\vec{x} = \vec{b}$ y, simultáneamente, encontrar la solución del sistema homogéneo asociado $A\vec{x} = \vec{0}$ escribiendo la solución del sistema $A\vec{x} = \vec{b}$ en la forma $\vec{p} + \vec{h}$, donde $\vec{p}$ es una solución particular del sistema no homogéneo y $\vec{h}$ es la solución del sistema homogéneo asociado.

96 $$\left[\begin{array}{ccccc|c} 1 & -1 & 2 & 3 & 4 & 4 \end{array}\right]$$

97 $$\left[\begin{array}{cccc|c} 4 & -2 & 2 & 0 & -6 \\ 4 & 2 & 0 & -1 & 1 \end{array}\right]$$

98 $$\left[\begin{array}{cccc|c} 1 & -2 & 1 & 1 & 5 \\ 2 & 1 & -3 & -1 & -4 \\ 1 & -7 & -6 & 2 & -13 \end{array}\right]$$

99 $$\left[\begin{array}{cccc|c} 2 & 1 & 3 & 0 & 4 \\ 1 & -1 & 2 & 1 & 1 \\ 4 & -1 & 7 & 2 & 6 \\ -1 & -2 & -1 & 1 & -3 \end{array}\right]$$

100 $$\left[\begin{array}{ccccc|c} 1 & 2 & -1 & 2 & 1 & -8 \\ 2 & 4 & 1 & -2 & 3 & -9 \\ 3 & 6 & 2 & -6 & 5 & -11 \end{array}\right]$$

101 $$\left[\begin{array}{cccc|c} 1 & 3 & -1 & 2 & 2 \\ 0 & 11 & -5 & 3 & 4 \\ 2 & -5 & 3 & 1 & 0 \\ 4 & 1 & 1 & 5 & 4 \end{array}\right]$$

102 Determinar los valores de α para que el sistema

$$\begin{aligned} x + 2y - 3z &= 4 \\ 3x - y + 5z &= 2 \\ 4x + y + (\alpha^2 - 14) &= a + 2 \end{aligned}$$

tenga solución única.

103 Encontrar los valores de α para los que el sistema del ejercicio 104 sea inconsistente.

104 ¿Para qué valores de α el sistema del ejercicio 104 tiene una infinidad de soluciones?

En los ejercicios 105 a 107 determinar los valores de α para que el sistema dado tenga *(a)* solución única, *(b)* sea inconsistente, *(c)* tenga una infinidad de soluciones.

105 $$\begin{aligned} x_1 + \alpha x_2 + x_3 &= 1 \\ \alpha x_1 + x_2 + x_3 &= 1 \\ 2x_1 + (\alpha+1)x_2 + (\alpha+1)x_3 &= 2 \end{aligned}$$

106 $$\begin{aligned} x_1 + x_2 + \alpha x_3 &= 2 \\ 3x_1 + 4x_2 + 2x_3 &= \alpha \\ x_1 + x_2 + 3x_3 &= \alpha - 1 \end{aligned}$$

107 $$\begin{aligned} x_1 - 3x_3 &= -3 \\ 2x_1 + \alpha x_2 - x_3 &= -2 \\ 2x_1 + 2x_2 + (\alpha - 3)x_3 &= -2 \end{aligned}$$

En los ejercicios 108 a 109 determinar las condiciones que deben cumplir los parámetros α, β y γ para que el sistema dado sea consistente.

108 $$\begin{aligned} x_1 - 2x_2 + 3x_3 &= \alpha \\ 2x_1 - x_2 + 2x_3 &= \beta \\ -x_1 + 5x_2 - 7x_3 &= \gamma \end{aligned}$$

109 $$\begin{aligned} x_1 - x_2 + x_3 &= \alpha \\ 2x_1 - x_2 - 2x_3 &= \beta \\ -3x_1 + x_2 + 2x_3 &= \gamma \end{aligned}$$

110 Encontrar los valores de α tales que el sistema homogéneo

$$\begin{aligned} x_1+3x_2+(1-\alpha)x_3 &= 0\\ x_1+\alpha x_2-x_3 &= 0\\ (2-\alpha)x_1+3x_2 &= 0 \end{aligned}$$

tenga solución no trivial.

111 Resolver el sistema

$$\begin{aligned} z_1-iz_2+z_3+iz_3+4iz_4 &= 5+2i\\ -2z_1+2z_2+3z_3-2iz_4 &= 1-2i\\ 3iz_1-z_2+2z_3+z_4-iz_4 &= -2-i \end{aligned}$$

112 Encontrar las soluciones del sistema

$$\begin{aligned} z_1+iz_2-iz_3 &= 3\\ 2iz_1+z_2-iz_3 &= 1+i\\ -3z_1+iz_2-iz_3 &= -1 \end{aligned}$$

113 Resolver el sistema

$$\begin{aligned} z_1-z_2+z_3 &= -i\\ 2z_1-z_2+z_3 &= -2i\\ -2z_1+z_2+z_3 &= 4i \end{aligned}$$

114 Hallar las soluciones del sistema

$$\begin{aligned} z_1-z_2+z_3+z_4 &= 1+i\\ iz_1-z_2+z_3+2z_4 &= 0\\ -2z_1+iz_2-z_3+z_4 &= -3i\\ -z_1-z_2+iz_3+2z_4 &= 1-2i \end{aligned}$$

2 Matrices invertibles y determinantes

En este capítulo trataremos dos importantes conceptos de la teoría de matrices cuadradas. El primero involucra el estudio de cierto tipo de matrices que, análogamente a los números reales distintos de cero, poseen un inverso multiplicativo; el segundo, estrechamente relacionado con el primero, versa sobre un número llamado determinante que se le asocia a cada matriz cuadrada y que tiene importantes y variadas aplicaciones en álgebra lineal. Al finalizar estos temas se incluye una sección de ejercicios resueltos y propuestos.

2.1 Matrices invertibles y sus inversas

Ya vimos que las operaciones algebraicas con matrices tienen propiedades análogas a las de los números reales; en esta sección daremos una más, que corresponde a la existencia de inversos multiplicativos.

2.1.1 Definición y propiedades

Definición 2.1 *Una matriz cuadrada A de orden n es* **invertible**,[1] *si existe una matriz C del mismo tamaño tal que*

$$AC = CA = I_n.$$

► **Ejemplo 2.1** Si $A = \begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix}$ y $C = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix}$, la matriz A es invertible, pues,

$$AC = \begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

y

$$CA = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. ◄$$

El teorema siguiente garantiza que, de existir, la matriz C de la definición 2.1 es única.

[1] Si una matriz es invertible, se acostumbra decir que es **no singular**; y en caso contrario que es **singular**.

Teorema 2.1 *Si $A \in \mathfrak{M}_{n\times n}$ es invertible y $C, B \in \mathfrak{M}_{n\times n}$ son tales que*

$$AC = CA = I_n \quad y \quad AB = BA = I_n,$$

entonces

$$B = C.$$

DEMOSTRACIÓN ■ En efecto, $AB = I_n \Rightarrow C(AB) = CI_n \Rightarrow (CA)B = C \Rightarrow I_nB = C \Rightarrow B = C.$ ■

Definición 2.2 *Si $A \in \mathfrak{M}_{n\times n}$ es invertible, a la única matriz $C \in \mathfrak{M}_{n\times n}$ tal que $AC = CA = I_n$ se le llama la* **matriz inversa**[2] *de A y se le denota como A^{-1}; es decir, $C = A^{-1}$.*

► **Ejemplo 2.2** Del ejemplo precedente

$$\begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix}. ◄$$

❍ **Nota 2.1** De la demostración del teorema 2.1 se desprende que basta probar $AC = I_n$ o $CA = I_n$ para que $C = A^{-1}$.

Teorema 2.2 ***(Propiedades)***

1. *Si $A, B \in \mathfrak{M}_{n\times n}$ son matrices invertibles, entonces la matriz AB es invertible y además $(AB)^{-1} = B^{-1}A^{-1}$.*
2. *Si $A \in \mathfrak{M}_{n\times n}$ es una matriz invertible, A^{-1} es también invertible y $(A^{-1})^{-1} = A$.*
3. *Si A es una matriz cuadrada invertible, entonces A^t es también invertible y además $(A^t)^{-1} = (A^{-1})^t$.*
4. *Si $\alpha \neq 0$ y A es una matriz cuadrada invertible, entonces αA es invertible y $(\alpha A)^{-1} = (1/\alpha)A^{-1}$.*

DEMOSTRACIÓN ■ 1.
$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A((BB^{-1})A^{-1}) \\ &= A(I_nA^{-1}) \\ &= AA^{-1} \\ &= I_n.\end{aligned}$$

Por tanto, $(AB)^{-1} = B^{-1}A^{-1}$.

2. $A^{-1}A = I_n$, por tanto, $(A^{-1})^{-1} = A$.
3. Puesto que (cfr. propiedad 3*(b)* en la pág. 10) $(MN)^t = N^tM^t$, las siguientes implicaciones son válidas.

$$\begin{aligned}(AA^{-1})^t = I_n^t = I_n \quad &\Rightarrow \quad (A^{-1})^tA^t = I_n \\ &\Rightarrow \quad (A^t)^{-1} = (A^{-1})^t.\end{aligned}$$

4. Se deja como ejercicio al lector. ■

[2] El lector debe tener siempre presente que los conceptos de matriz invertible y de su inversa sólo se definen en matrices cuadradas.

2.1.2 Matrices invertibles y sistemas lineales

Supongamos que la matriz cuadrada $A \in \mathfrak{M}_{n\times n}$ es invertible, entonces el sistema cuadrado $A\vec{x}=\vec{b}$ tiene solución única para todo $\vec{b}$. En efecto,

$$A\vec{x}=\vec{b} \Rightarrow A^{-1}(A\vec{x}) = A^{-1}\vec{b} \Rightarrow \left(A^{-1}A\right)\vec{x} = A^{-1}\vec{b} \Rightarrow I_n\vec{x} = A^{-1}\vec{b};$$

por tanto

$$\vec{x} = A^{-1}\vec{b}.$$

▶ **Ejemplo 2.3** Utilizar el ejemplo 2.1 para resolver el sistema

$$\begin{aligned} 2x_1+9x_2 &= 4 \\ x_1+4x_2 &= -5 . \end{aligned}$$ ◀

Solución Si A es la matriz de coeficientes del sistema, por el mencionado ejemplo (cfr. pág. 63) A es invertible y

$$A^{-1} = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix},$$

entonces

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 4 \\ -5 \end{bmatrix} = \begin{bmatrix} -61 \\ 14 \end{bmatrix}.$$ ✔

En el ejemplo 2.1 mostramos que la matriz

$$C = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix} \quad \text{es la inversa de la matriz} \quad A = \begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix},$$

pero no indicamos cómo encontramos $C = A^{-1}$. Resolvamos esta cuestión en el siguiente ejemplo.

▶ **Ejemplo 2.4** Encontrar la inversa de la matriz A del ejemplo 2.1; es decir, hallar

$$\begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix}^{-1}.$$ ◀

Solución Busquemos una matriz $C = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$, tal que $AC = I_2$; esto es,

$$\begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x & z \\ y & w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Como

$$\begin{bmatrix} 2 & 9 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} 2x+9z & 2y+9w \\ x+4z & y+4w \end{bmatrix},$$

se debe cumplir

$$\begin{bmatrix} 2x+9z & 2y+9w \\ x+4z & y+4w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix};$$

por tanto,

$$\begin{aligned} 2x+9z &= 1 \\ x+4z &= 0 \end{aligned} \tag{2.1}$$

y

$$\begin{aligned} 2y+9w &= 0 \\ y+4w &= 1. \end{aligned} \tag{2.2}$$

Ambos sistemas tienen la misma matriz de coeficientes. Trabajemos simultáneamente los dos sistemas como lo hicimos en el capítulo anterior (cfr. pág. 27).

$$\begin{aligned} \left[\begin{array}{cc|cc} 2 & 9 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{array}\right] &\leftrightarrow \left[\begin{array}{cc|cc} 1 & 4 & 0 & 1 \\ 2 & 9 & 1 & 0 \end{array}\right] \\ &\leftrightarrow \left[\begin{array}{cc|cc} 1 & 4 & 0 & 1 \\ 0 & 1 & 1 & -2 \end{array}\right] \end{aligned} \tag{2.3}$$

Resolviendo por sustitución regresiva para el sistema (2.1), encontramos $z = 1$ y $x = -4$; mientras que para el sistema (2.2) tenemos $w = -2$ y $y = 1-4w = 9$; con lo que la matriz buscada es

$$C = \begin{bmatrix} -4 & 9 \\ 1 & -2 \end{bmatrix} \tag{2.4}$$

que es la matriz propuesta en el ejemplo 2.1. Apliquemos ahora el método de Gauss-Jordan a la matriz ampliada (2.3) para resolver los sistemas (2.1) y (2.2).

$$\begin{aligned} \left[\begin{array}{cc|cc} 2 & 9 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{array}\right] &\leftrightarrow \left[\begin{array}{cc|cc} 1 & 4 & 0 & 1 \\ 2 & 9 & 1 & 0 \end{array}\right] \\ &\leftrightarrow \left[\begin{array}{cc|cc} 1 & 4 & 0 & 1 \\ 0 & 1 & 1 & -2 \end{array}\right] \\ &\leftrightarrow \left[\begin{array}{cc|cc} 1 & 0 & -4 & 9 \\ 0 & 1 & 1 & -2 \end{array}\right] \end{aligned} \tag{2.5}$$

Donde nuevamente encontramos $z = 1$, $x = -4$, $y = 9$ y $w = -2$; es decir, la matriz C de (2.4); pero ésta es precisamente la matriz que está en el lado derecho de la partición en (2.5). ✔

Ya probamos que si $A \in \mathfrak{M}_{n\times n}$ es una matriz cuadrada invertible, entonces el sistema lineal $A\vec{x} = \vec{b}$ tiene solución única $\forall \vec{b}$ (cfr. pág. 65). Recíprocamente, supongamos que el sistema $A\vec{x} = \vec{b}$ tiene solución única $\forall \vec{b}$; en particular para cada $\vec{b} = \vec{e}_k = [0 \; \cdots \; 0 \; \overset{k}{1} \; 0 \; \cdots \; 0]^t$, $k = 1, 2, \ldots, n$. Sea, para cada $k = 1, \ldots, n$, $\vec{c}_k = [c_{1k} \; c_{2k} \; \cdots \; c_{nk}]^t$, la única solución del sistema $A\vec{x} = \vec{e}_k$. Por tanto,

$$A\vec{c}_k = \vec{e}_k$$

para cada $k = 1, \ldots, n$. Sea la matriz C cuyas columnas son las soluciones $\vec{c}_k$; es decir, $C = [c_{ik}]$, $1 \le i \le n$, $1 \le k \le n$. Entonces, por lo probado en el ejemplo 1.14 (cfr. (1.2) pág. 11), tenemos

$$AC = [A\vec{c}_1 \; A\vec{c}_2 \; \cdots \; A\vec{c}_n] = [\vec{e}_1 \; \vec{e}_2 \; \cdots \; \vec{e}_n];$$

es decir,

$$AC = I_n$$

y, por ende, la matriz A es invertible. Finalmente, del inciso 3 de los criterios para solución única de la página 31, sabemos que el sistema $A\vec{x} = \vec{b}$ tiene solución única para todo $\vec{b}$ si y sólo si la matriz A es equivalente a la identidad I_n. Hemos probado así el siguiente teorema.

Teorema 2.3 *Si $A \in \mathfrak{M}_{n\times n}$, las siguientes condiciones son equivalentes a pares:*[3]

1. *El sistema $A\vec{x} = \vec{b}$ tiene solución única $\forall \vec{b}$.*
2. *$A \sim I_n$.*
3. *A es invertible.*

► **Ejemplo 2.5** Hallar los valores de r tales que la matriz

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & r & 3 \\ 2 & 1 & 3 \end{bmatrix}$$

sea invertible.◄

Solución Por el teorema anterior, es suficiente que la matriz A sea equivalente a la identidad para que sea invertible. Entonces, si llevamos A a forma escalonada y r es tal que en cada columna hay pivote,[4] por el método de Gauss-Jordan A será, efectivamente, equivalente a la identidad. Llevemos entonces A a forma escalonada.

$$\begin{bmatrix} 1 & 2 & 0 \\ 1 & r & 3 \\ 2 & 1 & 3 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & r-2 & 3 \\ 0 & -3 & 3 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 0 & r-2 & 3 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & r+1 \end{bmatrix}.$$

Así, para cualquier valor $r \neq -1$, la matriz A será invertible. ✔

► **Ejemplo 2.6** Determine cuáles de las siguientes matrices son invertibles.

1. $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$.
2. $B = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 0 & 2 \\ 2 & -2 & 3 \end{bmatrix}$.◄

[3] Que las condiciones sean equivalentes a pares significa que si se cumple una, cualquiera de ellas, se cumplen también todas las demás; esto es: (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4).

[4] Véase la nota 1.6, página 31.

Solución 1. La matriz $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ está en forma escalonada y no tiene pivote en la segunda columna, por tanto no es equivalente a I_2, por el teorema 2.3 no es una matriz invertible.

2. Llevemos la matriz B a forma escalonada mediante el método de Gauss.

$$\begin{bmatrix} 1 & 1 & -1 \\ -1 & 0 & 2 \\ 2 & -2 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & -4 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 9 \end{bmatrix}.$$

Dado que todas las columnas en la forma escalonada tienen pivote, la matriz es equivalente a la identidad I_3, por el teorema 2.3 la matriz B es invertible. ✔

2.1.3 Método de Gauss-Jordan para hallar la inversa de una matriz

Podemos generalizar el procedimiento empleado para encontrar la inversa de una matriz dado en el ejemplo 2.4 (pág. 65). Sea A una matriz cuadrada de orden n y busquemos una matriz $X = [x_{ij}]$ tal que $AX = I_n$. Lo cual equivale a resolver los sistemas lineales con la misma matriz de coeficientes $A\vec{x}_k = \vec{e}_k$, donde $\vec{x}_k = [x_{1k}\ x_{2k}\ \cdots\ x_{nk}]^t$ y $\vec{e}_k = [0\ \cdots\ 0\ \overset{k}{1}\ 0\ \cdots\ 0]^t$ es la k-ésima columna de la identidad I_n, $k = 1, 2, \ldots, n$; pues (cfr. ejemplo 1.14, pág. 11)

$$AX = [A\vec{x}_1\ A\vec{x}_2\ \cdots\ A\vec{x}_n] = I_n \Leftrightarrow A\vec{x}_k = \vec{e}_k\,.$$

Recordemos (cfr. pág. 27) que para resolver los sistemas $[A\,|\,\vec{e}_k]$ formamos la matriz ampliada $[A\,|\,\vec{e}_1\ \vec{e}_2\ \ldots\ \vec{e}_n] = [A\,|\,I_n]$ y aplicamos el método de Gauss. Pero esta vez utilicemos mejor el método de Gauss-Jordan como lo hicimos en el ejemplo 2.4. Si a la mitad del proceso, cuando $[A\,|\,I_n]$ se ha llevado a una forma escalonada equivalente $[H\,|\,J]$, todas las columnas de H tienen pivote, entonces todos los sistemas $[A\,|\,\vec{e}_k]$ tendrán solución única y A será equivalente a la identidad; al continuar con el proceso de Gauss-Jordan eventualmente se llegará a la equivalencia $[A\,|\,I_n] \sim [I_n\,|\,B]$, donde la k-ésima columna de B será la única solución del sistema $[A\,|\,\vec{e}_k]$; es decir, B es la solución de la ecuación matricial $AX = I_n$. Por lo que la matriz A será invertible y $B = A^{-1}$. Si a la mitad del proceso, cuando $[A\,|\,I_n]$ se ha llevado a una forma escalonada $[H\,|\,J]$, existe una columna de H que no tiene pivote, entonces $A \nsim I_n$ y por ende A no será invertible o, equivalentemente, alguno de los sistemas $[A\,|\,\vec{e}_k]$ será inconsistente y entonces $AX = I_n$ no tendrá solución. Resumimos la información precedente a continuación.

Método de Gauss-Jordan para hallar la inversa de una matriz

Para determinar si una matriz A es invertible y hallar su inversa, en tal caso, se procede de la siguiente manera:

1. Formamos la matriz aumentada $[A\,|\,I]$. Donde I es la identidad del mismo tamaño de A.
2. Se lleva A mediante el método de Gauss-Jordan a la identidad, aplicando simultáneamente las mismas operaciones de renglón en el lado derecho de la partición.
3. Al llegar a $[I\,|\,B]$, $B = A^{-1}$.
4. Si A no se puede llevar a I, significa que A no es invertible.[5]

[5] Recordemos que para esto basta que en una forma escalonada equivalente, obtenida con el método de Gauss, no haya pivote en alguna columna.

▶ **Ejemplo 2.7** Hallar, si es que existe, la inversa de la matriz

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}. \blacktriangleleft$$

Solución

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right] \xleftrightarrow[R_3 \leftrightarrow -R_1 + R_3]{R_2 \leftrightarrow -R_1 + R_2} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right]$$

$$\xleftrightarrow{R_3 \leftrightarrow -2R_2 - R_3} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

$$\xleftrightarrow[R_1 \leftrightarrow -3R_3 + R_1]{R_2 \leftrightarrow 3R_3 + R_2} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

$$\xleftrightarrow{R_1 \leftrightarrow -2R_2 + R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right].$$

Por tanto, tenemos

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}.$$

Comprobación:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix} \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad ✔$$

▶ **Ejemplo 2.8** Encontrar, si es que existe, la inversa de la matriz

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & -3 \\ -1 & -2 & -6 \end{bmatrix}. \blacktriangleleft$$

Solución

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & -3 & 0 & 1 & 0 \\ -1 & -2 & -6 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -5 & -2 & 1 & 0 \\ 0 & -1 & -5 & 1 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -5 & -2 & 1 & 0 \\ 0 & 0 & 0 & 3 & -1 & 1 \end{array}\right] \tag{2.6}$$

La forma escalonada (2.6) equivalente a $[A \mid I_3]$ no tiene pivote en la tercera columna a la izquierda de la partición; por tanto la matriz A no es invertible (pues no es equivalente a I_3). ✔

▶ **Ejemplo 2.9** Hallar, si es que existe, la inversa de la matriz

$$A=\begin{bmatrix} 1 & -2 & 1 & 0\\ -3 & 5 & 0 & 2\\ 0 & 1 & 2 & -4\\ -1 & 2 & 4 & -1\end{bmatrix}.◀$$

Solución

$$\left[\begin{array}{rrrr|rrrr} 1 & -2 & 1 & 0 & 1 & 0 & 0 & 0\\ -3 & 5 & 0 & 2 & 0 & 1 & 0 & 0\\ 0 & 1 & 2 & -4 & 0 & 0 & 1 & 0\\ -1 & 2 & 4 & -1 & 0 & 0 & 0 & 1\end{array}\right] \sim \left[\begin{array}{rrrr|rrrr} 1 & -2 & 1 & 0 & 1 & 0 & 0 & 0\\ 0 & -1 & 3 & 2 & 3 & 1 & 0 & 0\\ 0 & 1 & 2 & -4 & 0 & 0 & 1 & 0\\ 0 & 0 & 5 & -1 & 1 & 0 & 0 & 1\end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|rrrr} 1 & -2 & 1 & 0 & 1 & 0 & 0 & 0\\ 0 & -1 & 3 & 2 & 3 & 1 & 0 & 0\\ 0 & 0 & 5 & -2 & 3 & 1 & 1 & 0\\ 0 & 0 & 5 & -1 & 1 & 0 & 0 & 1\end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|rrrr} 1 & -2 & 1 & 0 & 1 & 0 & 0 & 0\\ 0 & -1 & 3 & 2 & 3 & 1 & 0 & 0\\ 0 & 0 & 5 & -2 & 3 & 1 & 1 & 0\\ 0 & 0 & 0 & 1 & -2 & -1 & -1 & 1\end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|rrrr} 1 & -2 & 1 & 0 & 1 & 0 & 0 & 0\\ 0 & -1 & 3 & 0 & 7 & 3 & 2 & -2\\ 0 & 0 & 5 & 0 & -1 & -1 & -1 & 2\\ 0 & 0 & 0 & 1 & -2 & -1 & -1 & 1\end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|rrrr} 1 & -2 & 1 & 0 & 1 & 0 & 0 & 0\\ 0 & -1 & 3 & 0 & 7 & 3 & 2 & -2\\ 0 & 0 & 1 & 0 & -1/5 & -1/5 & -1/5 & 2/5\\ 0 & 0 & 0 & 1 & -2 & -1 & -1 & 1\end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|rrrr} 1 & -2 & 0 & 0 & 6/5 & 1/5 & 1/5 & -2/5\\ 0 & -1 & 0 & 0 & 38/5 & 18/5 & 13/5 & -16/5\\ 0 & 0 & 1 & 0 & -1/5 & -1/5 & -1/5 & 2/5\\ 0 & 0 & 0 & 1 & -2 & -1 & -1 & 1\end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|rrrr} 1 & 0 & 1 & 0 & -14 & -7 & -5 & 6\\ 0 & -1 & 0 & 0 & 38/5 & 18/5 & 13/5 & -16/5\\ 0 & 0 & 1 & 0 & -1/5 & -1/5 & -1/5 & 2/5\\ 0 & 0 & 0 & 1 & -2 & -1 & -1 & 1\end{array}\right]$$

$$\sim \left[\begin{array}{rrrr|rrrr} 1 & 0 & 1 & 0 & -14 & -7 & -5 & 6\\ 0 & 1 & 0 & 0 & -38/5 & -18/5 & -13/5 & 16/5\\ 0 & 0 & 1 & 0 & -1/5 & -1/5 & -1/5 & 2/5\\ 0 & 0 & 0 & 1 & -2 & -1 & -1 & 1\end{array}\right].$$

De donde A es invertible y

$$A^{-1}=\begin{bmatrix} -14 & -7 & -5 & 6\\ -38/5 & -18/5 & -13/5 & 16/5\\ -1/5 & -1/5 & -1/5 & 2/5\\ -2 & -1 & -1 & 1\end{bmatrix}.$$

Comprobando:

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ -3 & 5 & 0 & 2 \\ 0 & 1 & 2 & -4 \\ -1 & 2 & 4 & -1 \end{bmatrix} \begin{bmatrix} -14 & -7 & -5 & 6 \\ -38/5 & -18/5 & -13/5 & 16/5 \\ -1/5 & -1/5 & -1/5 & 2/5 \\ -2 & -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad ✔$$

2.1.4 Matrices elementales

Definición 2.3 *Una matriz cuadrada de orden m es* **elemental** *si se obtiene a partir de la identidad I_m mediante una sola operación de renglón.*

▶ **Ejemplo 2.10** La matriz $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ es elemental, pues se obtiene de I_3 mediante la operación de renglón $R_2 \leftrightarrow 2R_2$. Análogamente las matrices

$$E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

son elementales; pues resultan de I_3 al aplicarle las operaciones $R_1 \leftrightarrow R_2$ y $R_3 \leftrightarrow -R_1 + R_3$, respectivamente (¡chéquelo!).◀

▶ **Ejemplo 2.11** Si $A = \begin{bmatrix} 2 & -3 & 0 & -1 \\ 2 & -1 & 0 & 2 \\ 0 & -1 & 2 & 3 \end{bmatrix}$, sea B la matriz que se obtiene de A al aplicarle la operación de renglón $R_2 \leftrightarrow R_3$ y sea E_1 la matriz elemental que resulta de la identidad I_3 al aplicarle la misma operación; es decir,

$$B = \begin{bmatrix} 2 & -3 & 0 & -1 \\ 0 & -1 & 2 & 3 \\ 2 & -1 & 0 & 2 \end{bmatrix} \quad \text{y} \quad E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Entonces,

$$E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & -3 & 0 & -1 \\ 2 & -1 & 0 & 2 \\ 0 & -1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 0 & -1 \\ 0 & -1 & 2 & 3 \\ 2 & -1 & 0 & 2 \end{bmatrix} = B. ◀$$

Del ejemplo anterior deducimos el siguiente teorema, que no es difícil de probar y cuya demostración se deja como ejercicio al lector.

Teorema 2.4 ***(Efecto de matrices elementales)*** *Si a una matriz A de tamaño $m \times n$ se le aplica una operación de renglón, la matriz resultante es el producto EA, donde E es la matriz elemental $m \times m$ que se obtiene de la identidad I_m al aplicarle la misma operación de renglón.*

Supongamos ahora que E_1, E_2, E_3 son las matrices elementales $n\times n$ que se obtienen de la identidad al aplicarle, respectivamente, las operaciones de renglón $R_i \leftrightarrow R_j$, $R_i \leftrightarrow \alpha R_i$, $R_i \leftrightarrow \alpha R_i + \beta R_j$. Entonces, si E_1', E_2' y E_3' son las matrices elementales que resultan de la identidad al aplicarle las operaciones de fila $R_j \leftrightarrow R_i$, $R_i \leftrightarrow (1/\alpha)R_i$ y $R_i \leftrightarrow (1/\alpha)R_i - (\beta/\alpha)R_j$, respectivamente, es claro que al aplicar a E_1, E_2, E_3 sendas operaciones se obtendrá, en cada caso, nuevamente la identidad (¡compruébelo!). Así, por el efecto que tienen las matrices elementales se tiene: $E_1'E_1 = I_n$, $E_2'E_2 = I_n$ y $E_3'E_3 = I_n$. Lo cual demuestra la siguiente afirmación.

Teorema 2.5 *Toda matriz elemental es invertible y su inversa es también una matriz elemental.*

► **Ejemplo 2.12** Sean las matrices elementales

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & -3 \end{bmatrix};$$

las cuales se obtienen de I_3 mediante las operaciones $R_2 \leftrightarrow R_3$, $R_2 \leftrightarrow -3R_2$, $R_3 \leftrightarrow -3R_3 + 2R_1$, respectivamente. Entonces, para este caso

$$E_1' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad E_2' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2/3 & 0 & -1/3 \end{bmatrix} \quad \text{y}$$

$$E_1'E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$E_2'E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$E_3'E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2/3 & 0 & -1/3 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. ◄$$

► **Ejemplo 2.13** Hallar una matriz C, si es que existe, tal que la matriz CA sea triangular superior si

$$A = \begin{bmatrix} 0 & 1 & -3 \\ 2 & 3 & -1 \\ 4 & 5 & -2 \end{bmatrix}. ◄$$

Solución Llevemos A, mediante el método de Gauss, a forma triangular superior:

$$\begin{bmatrix} 0 & 1 & -3 \\ 2 & 3 & -1 \\ 4 & 5 & -2 \end{bmatrix} \leftrightarrow \begin{bmatrix} 4 & 5 & -2 \\ 2 & 3 & -1 \\ 0 & 1 & -3 \end{bmatrix} \leftrightarrow \begin{bmatrix} 4 & 5 & -2 \\ 0 & -1 & 0 \\ 0 & 1 & -3 \end{bmatrix} \leftrightarrow \begin{bmatrix} 4 & 5 & -2 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

Las matrices elementales correspondientes a la sucesión de estas operaciones son, respectivamente,

$$E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Luego,

$$E_3(E_2(E_1A)) = \begin{bmatrix} 4 & 5 & -2 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

Y como $E_3(E_2(E_1A)) = (E_3E_2E_1)A$, si hacemos $C = E_3E_2E_1$; es decir,

$$\begin{aligned} C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 1 & -2 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & -2 & 1 \\ 1 & -2 & 1 \end{bmatrix}, \end{aligned}$$

tenemos que

$$CA = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -2 & 1 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & -3 \\ 2 & 3 & -1 \\ 4 & 5 & -2 \end{bmatrix} = \begin{bmatrix} 4 & 5 & -2 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

es una matriz triangular superior. ✔

Ya vimos que un sistema cuadrado $A\vec{x} = \vec{b}$ tiene solución única $\forall\vec{b}$ si y sólo si la matriz A es equivalente a la identidad I. Ahora supongamos que efectivamente A es equivalente a I, entonces podemos llevar A mediante operaciones de renglón a la identidad; luego, por el efecto que tienen las matrices elementales, existen $E_k, E_{k-1}, \ldots, E_1$, matrices elementales del mismo orden de A, tales que $(E_kE_{k-1}\cdots E_1)A = I$. Es decir, la matriz A es invertible con inversa $E_kE_{k-1}\cdots E_1$. Por otro lado, de la igualdad $(E_kE_{k-1}\cdots E_1)A = I$ se desprende

$$A = E_1^{-1}\cdots E_{k-1}^{-1}E_k^{-1}.$$

Lo cual nos dice que A es producto de matrices elementales (pues la inversa de una matriz elemental es también elemental). Recíprocamente si A es producto de matrices elementales, entonces A es invertible. De lo precedente podemos añadir una condición más de equivalencia al teorema 2.3 (cfr. pág. 67).

Teorema 2.6 *Si $A \in \mathfrak{M}_{n\times n}$, las siguientes condiciones son equivalentes a pares:*

1. *El sistema $A\vec{x} = \vec{b}$ tiene solución única $\forall\vec{b}$.*
2. *$A \sim I_n$.*
3. *A es invertible.*
4. *A es producto de matrices elementales.*

2.1.5 Inversas de matrices con componentes complejas

La identidad multiplicativa en el sistema de números complejos[6] coincide con la identidad multiplicativa en los números reales, el número $z = 1$. Esto implica que la matriz identidad en $\mathfrak{M}_n(\mathbb{C})$, el conjunto de matrices cuadradas de orden n con componentes complejas, es la matriz identidad I_n con la que hemos trabajado en matrices con entradas reales; es decir, $I_n = [a_{kl}]$, donde

$$a_{kl} = \begin{cases} 1 & \text{si } k = l \\ 0 & \text{si } k \neq l \end{cases}$$

y satisface (como en el caso real)

$$AI_n = A$$
$$I_nB = B$$

para toda $A \in \mathfrak{M}_{m\times n}(\mathbb{C})$ y para toda $B \in \mathfrak{M}_{n\times m}(\mathbb{C})$. La definición de matriz invertible es exactamente la misma para $\mathfrak{M}_n(\mathbb{C})$ que para el caso real (cfr. definición 2.1, pág. 63). Nuevamente todas las propiedades y teoremas acerca de las matrices invertibles y sus inversas dados en esta sección son también válidos para el caso de componentes complejas. Entonces, para determinar si una matriz en $\mathfrak{M}_n(\mathbb{C})$ es invertible y en tal caso calcular su inversa, se utilizan las mismas técnicas que en el caso real pero se trabaja con el álgebra de números complejos.

► **Ejemplo 2.14** Determinar si la matriz

$$A = \begin{bmatrix} 1 & 1+i & -i \\ i & -2i & 1 \\ 1 & 1-i & 2i \end{bmatrix}$$

es invertible. De ser así, hallar A^{-1}.◄

Solución Apliquemos el método de Gauss-Jordan.[7]

$$\left[\begin{array}{ccc|ccc} 1 & 1+i & -i & 1 & 0 & 0 \\ i & -2i & 1 & 0 & 1 & 0 \\ 1 & 1-i & 2i & 0 & 0 & 1 \end{array}\right] \overset{(1)}{\longleftrightarrow} \left[\begin{array}{ccc|ccc} 1 & 1+i & -i & 1 & 0 & 0 \\ 0 & 1-3i & 0 & -i & 1 & 0 \\ 0 & -2i & 3i & -1 & 0 & 1 \end{array}\right]$$

$$\overset{(2)}{\longleftrightarrow} \left[\begin{array}{ccc|ccc} 1 & 1+i & -i & 1 & 0 & 0 \\ 0 & -2i & 3i & -1 & 0 & 1 \\ 0 & 1-3i & 0 & -i & 1 & 0 \end{array}\right]$$

$$\overset{(3)}{\longleftrightarrow} \left[\begin{array}{ccc|ccc} 1 & 1+i & -i & 1 & 0 & 0 \\ 0 & 1 & -3/2 & -i/2 & 0 & i/2 \\ 0 & 1-3i & 0 & -i & 1 & 0 \end{array}\right]$$

[6] Cfr. B.1 del apéndice B.

[7] (1) $R_2 \leftrightarrow -iR_1 + R_2$, $R_3 \leftrightarrow -R_1 + R_2$, (2) $R_2 \leftrightarrow R_3$, (3) $R_2 \leftrightarrow (-1/2i)R_2$, (4) $R_3 \leftrightarrow -(1-3i)R_2 + R_3$, (5) $R_3 \leftrightarrow (3/2-(9/2)i)^{-1}R_3$, (6) $R_2 \leftrightarrow (3/2)R_3 + R_2$, $R_1 \leftrightarrow iR_3 + R_1$, (7) $R_1 \leftrightarrow -(1+i)R_2 + R_1$.

$$\overset{(4)}{\longleftrightarrow}\left[\begin{array}{ccc|ccc} 1 & 1+i & -i & 1 & 0 & 0 \\ 0 & 1 & -3/2 & -i/2 & 0 & i/2 \\ 0 & 0 & 3/2-9i/2 & \frac{3}{2}-\frac{i}{2} & 1 & -\frac{3}{2}-\frac{i}{2} \end{array}\right]$$

$$\overset{(5)}{\longleftrightarrow}\left[\begin{array}{ccc|ccc} 1 & 1+i & -i & 1 & 0 & 0 \\ 0 & 1 & -3/2 & -i/2 & 0 & i/2 \\ 0 & 0 & 1 & \frac{1}{5}+\frac{4}{15}i & \frac{1}{15}+\frac{1}{5}i & -\frac{1}{3}i \end{array}\right]$$

$$\overset{(6)}{\longleftrightarrow}\left[\begin{array}{ccc|ccc} 1 & 1+i & 0 & \frac{11}{15}+\frac{1}{5}i & -\frac{1}{5}+\frac{1}{15}i & \frac{1}{3} \\ 0 & 1 & 0 & \frac{3}{10}-\frac{1}{10}i & \frac{1}{10}+\frac{3}{10}i & 0 \\ 0 & 0 & 1 & \frac{1}{5}+\frac{4}{15}i & \frac{1}{15}+\frac{1}{5}i & -\frac{1}{3}i \end{array}\right]$$

$$\overset{(7)}{\longleftrightarrow}\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & -\frac{1}{3}i & \frac{1}{3} \\ 0 & 1 & 0 & \frac{3}{10}-\frac{1}{10}i & \frac{1}{10}+\frac{3}{10}i & 0 \\ 0 & 0 & 1 & \frac{1}{5}+\frac{4}{15}i & \frac{1}{15}+\frac{1}{5}i & -\frac{1}{3}i \end{array}\right].$$

Por tanto,

$$A^{-1}=\begin{bmatrix} \frac{1}{3} & -\frac{1}{3}i & \frac{1}{3} \\ \frac{3}{10}-\frac{1}{10}i & \frac{1}{10}+\frac{3}{10}i & 0 \\ \frac{1}{5}+\frac{4}{15}i & \frac{1}{15}+\frac{1}{5}i & -\frac{1}{3}i \end{bmatrix}. \checkmark$$

2.2 Determinantes

En este sección estudiaremos uno de los principales conceptos del álgebra lineal. A cada matriz cuadrada se le puede asignar cierto número real llamado el determinante de la misma; el cual tiene diversas e importantes aplicaciones en esta materia y otras áreas como, por citar algunos ejemplos, el cálculo del polinomio característico para encontrar los valores y vectores propios, método de la adjunta para hallar la inversa de una matriz, regla de Cramer, cálculo de volúmenes e integrales, etc. Utilizaremos un método de recurrencia para definir y calcular determinantes de orden n a partir de determinantes de orden menor. Hacemos énfasis en que, como en el caso de la inversa de una matriz, el concepto de determinante sólo es aplicable a matrices cuadradas.

2.2.1 Desarrollo por cofactores

Definición 2.4

1. *Si* $A=[a]$ *es una matriz* 1×1 *el* **determinante** *de* A *se denota y define como*

$$\det(A)=a.$$

2. *Si* $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$

es una matriz 2×2 se define y denota el determinante de A como

$$\det(A) = ad - bc.$$

También usaremos la notación

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

para el determinante de la matriz A.

3. *Supongamos que n es un entero mayor a 2 y que los determinantes para matrices de orden menor o igual a $n-1$ ya han sido definidos. Si A es una matriz $n\times n$, entonces:*
   (a) *Se define el* **menor** *del elemento a_{ij} de A (o simplemente el menor ij) como el determinante que se obtiene de la matriz que resulta de eliminar la fila i y la columna j de A. A este número lo denotamos por M_{ij}.*
   (b) *El* **cofactor** *del elemento a_{ij} de A (o simplemente el cofactor ij), se define y denota como*

$$c_{ij} = (-1)^{i+j} M_{ij}.$$

▶ **Ejemplo 2.15** Si

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -4 & 3 & 1 \\ 0 & 2 & 1 \end{bmatrix},$$

entonces

$$M_{12} = \det\begin{vmatrix} -4 & 1 \\ 0 & 1 \end{vmatrix} = -4$$

y

$$c_{12} = (-1)^{1+2} M_{12} = (-1)^3(-4) = 4. ◀$$

A continuación enunciamos el teorema básico para definir y calcular inductivamente los determinantes. Este teorema es conocido como **desarrollo de Laplace** o **desarrollo por cofactores** de la función determinante. La demostración de este teorema es larga y remitimos al lector al apéndice C para que la consulte. Únicamente nos limitamos, en los dos siguientes ejemplos, a ilustrar este importante teorema para los casos $n = 2$ y $n = 3$, e inmediatamente pasamos al cálculo de determinantes utilizando este resultado en los ejemplos subsecuentes.

Teorema 2.7 *Sea $A \in \mathfrak{M}_{n\times n}$.*

1. *Sean $f_1, f_2, \ldots, f_n$ los elementos de una fila, F_i, cualquiera de A, con $c_1, c_2, \ldots, c_n$ sus respectivos cofactores y*

$$\Delta_{F_i} = c_1 f_1 + c_2 f_2 + \cdots + c_n f_n.$$

2. *Sean $g_1, g_2, \ldots, g_n$ los elementos de una columna, G_j, cualquiera de A, con $d_1, d_2, \ldots, d_n$ sus respectivos cofactores y*

$$\Delta_{G_j} = d_1 g_1 + d_2 g_2 + \cdots + d_n g_n.$$

Entonces:

- $\Delta_{F_i} = \Delta_{F_k}$ *para todo par de filas* F_i *y* F_k *de* A.
- $\Delta_{G_j} = \Delta_{G_l}$ *para todo par de columnas* G_j *y* G_l *de* A.
- $\Delta_{F_i} = \Delta_{G_j}$ *para toda fila* F_i *y toda columna* G_j *de* A.

Definición 2.5 *Al número real común que se calcula en los incisos 1 y 2 del teorema anterior se le llama el* **determinante** *de la matriz* A *y se denota como* $\det(A)$ *o* $|A|$.[8] *Es decir,*

1. $|A| = \Delta_F$ *para cualquier fila* F *de* A *y*
2. $|A| = \Delta_G$ *para cualquier columna* G *de* A.

En el caso 1 se dice que el determinante se ha desarrollado por cofactores en la fila F *y, en el caso 2, que se ha desarrollado por cofactores en la columna* G.

▶ **Ejemplo 2.16** Sea $A = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$, entonces $c_{11} = (-1)^{1+1}M_{11} = \det[d] = d$; $c_{12} = (-1)^{1+2}M_{12} = -\det([c]) = -c$; $c_{21} = (-1)^{2+1}M_{21} = -\det([b]) = -b$; y $c_{22} = (-1)^{2+2}M_{22} = \det([a]) = a$. Por tanto,

$$\begin{aligned}
\Delta F_1 &= ac_{11} + bc_{12} = ad + b(-c) = ad - bc = \det(A), \\
\Delta F_2 &= cc_{21} + dc_{22} = c(-b) + da = ad - bc = \det(A), \\
\Delta G_1 &= ac_{11} + cc_{21} = ad + c(-b) = ad - bc = \det(A), \\
\Delta G_2 &= bc_{12} + dc_{22} = b(-c) + d(a) = ad - bc = \det(A). \blacktriangleleft
\end{aligned}$$

▶ **Ejemplo 2.17** Sea $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

Entonces

$$\begin{aligned}
\Delta F_1 &= a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13} \\
&= (-1)^{1+1}a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{1+2}a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^{1+3}a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\
&= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\
&= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31},
\end{aligned}$$

$$\begin{aligned}
\Delta F_2 &= a_{21}c_{21} + a_{22}c_{22} + a_{23}c_{23} \\
&= (-1)^{2+1}a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{2+2}a_{22}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^{2+3}a_{23}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \\
&= -a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) - a_{23}(a_{11}a_{32} - a_{12}a_{31}) \\
&= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31},
\end{aligned}$$

[8] El lector debe tener cuidado de no confundir esta notación con el valor absoluto de un número real. En realidad la notación $|A|$ quedará clara en contexto; pues si A es una matriz, entonces se refiere al determinante de la misma; mientras que si A es un número real representará, como es usual, su valor absoluto. Por esta razón, el lector debe tener siempre en mente que el determinante de una matriz puede ser un número real positivo, negativo o cero.

$$\begin{aligned}
\Delta F_3 &= a_{31}c_{31}+a_{32}c_{32}+a_{33}c_{33}\\
&= (-1)^{3+1}a_{31}\begin{vmatrix} a_{12} & a_{13}\\ a_{22} & a_{23}\end{vmatrix}+(-1)^{3+2}a_{32}\begin{vmatrix} a_{11} & a_{13}\\ a_{21} & a_{23}\end{vmatrix}+(-1)^{3+3}a_{33}\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}\\
&= a_{31}(a_{12}a_{23}-a_{13}a_{22})-a_{32}(a_{11}a_{23}-a_{13}a_{21})+a_{33}(a_{11}a_{22}-a_{12}a_{21})\\
&= a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31},
\end{aligned}$$

$$\begin{aligned}
\Delta G_1 &= a_{11}c_{11}+a_{21}c_{21}+a_{31}c_{31}\\
&= (-1)^{1+1}a_{11}\begin{vmatrix} a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}+(-1)^{2+1}a_{21}\begin{vmatrix} a_{12} & a_{13}\\ a_{32} & a_{33}\end{vmatrix}+(-1)^{3+1}a_{31}\begin{vmatrix} a_{12} & a_{13}\\ a_{22} & a_{23}\end{vmatrix}\\
&= a_{11}(a_{22}a_{33}-a_{23}a_{32})-a_{21}(a_{12}a_{33}-a_{13}a_{32})+a_{31}(a_{12}a_{23}-a_{13}a_{22})\\
&= a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31},
\end{aligned}$$

$$\begin{aligned}
\Delta G_2 &= a_{12}c_{21}+a_{22}c_{22}+a_{32}c_{23}\\
&= (-1)^{2+1}a_{12}\begin{vmatrix} a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix}+(-1)^{2+2}a_{22}\begin{vmatrix} a_{11} & a_{13}\\ a_{31} & a_{33}\end{vmatrix}+(-1)^{3+2}a_{32}\begin{vmatrix} a_{11} & a_{13}\\ a_{21} & a_{23}\end{vmatrix}\\
&= -a_{12}(a_{21}a_{33}-a_{23}a_{31})+a_{22}(a_{11}a_{33}-a_{13}a_{31})-a_{32}(a_{11}a_{23}-a_{13}a_{21})\\
&= a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31},
\end{aligned}$$

$$\begin{aligned}
\Delta G_3 &= a_{13}c_{13}+a_{23}c_{23}+a_{33}c_{33}\\
&= (-1)^{1+3}a_{13}\begin{vmatrix} a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}+(-1)^{2+3}a_{23}\begin{vmatrix} a_{11} & a_{12}\\ a_{31} & a_{32}\end{vmatrix}+(-1)^{3+3}a_{33}\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}\\
&= a_{13}(a_{21}a_{32}-a_{22}a_{31})-a_{23}(a_{11}a_{32}-a_{12}a_{31})+a_{33}(a_{11}a_{22}-a_{12}a_{21})\\
&= a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}.
\end{aligned}$$

De donde se ve que $\Delta F_i = \Delta F_k$ para todo par de filas de A, $\Delta G_j = \Delta G_l$ para todo par de columnas de A, y $\Delta F_i = \Delta G_j$ para cualquier fila y para cualquier columna de A.◄

► **Ejemplo 2.18**

$$\begin{aligned}
\begin{vmatrix} -1 & 0 & 2\\ 2 & 1 & 3\\ -2 & 1 & 1\end{vmatrix} &= 0\cdot c_{12}+c_{22}+c_{32}\\
&= (-1)^{2+2}\begin{vmatrix} -1 & 2\\ -2 & 1\end{vmatrix}+(-1)^{3+2}\begin{vmatrix} -1 & 2\\ 2 & 3\end{vmatrix}\\
&= -1-(-4)-(-3-4)=3-(-7)\\
&= 10.
\end{aligned}$$
◄

El determinante de una matriz 3×3, si se agrupan los términos de $\Delta F_i = \Delta G_j$ del ejemplo 2.17, puede escribirse como

$$\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} = \left\{\begin{array}{l}(a_{11}a_{22}a_{33}+a_{21}a_{32}a_{13}+a_{31}a_{23}a_{12})\\ -(a_{13}a_{22}a_{31}+a_{23}a_{32}a_{11}+a_{33}a_{21}a_{12})\end{array}\right\} \tag{2.7}$$

Ésta es la llamada regla de Sarrus y un ardid nemotécnico para tenerla siempre en mente es mediante el esquema contenido en la figura 2-1, que seguramente es familiar al lector de cursos en educación básica.

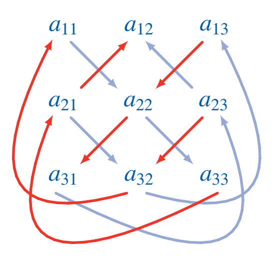


Figura 2-1 •

Los términos del primer sumando en la regla de Sarrus se obtienen multiplicando las componentes de la matriz siguiendo las flechas azules y los del segundo sumando multiplicando las componentes siguiendo las flechas rojas. Desafortunadamente esta regla sólo funciona para matrices 3×3 y su generalización ya no es válida para matrices de orden mayor que 3. El lector puede confirmar esta afirmación fácilmente si calcula un determinante 4×4 por la "generalización" de la regla de Sarrus y verifica el resultado desarrollando por cofactores.

Es posible hacer un poco más eficiente el desarrollo por cofactores. Si A es una matriz cuadrada de orden n y vamos a calcular su determinante, por ejemplo, desarrollando por cofactores en la fila F_i, entonces

$$|A| = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} M_{ik}$$

donde M_{ik} es el menor del elemento a_{ik} de la matriz A; es decir, M_{ik} es el determinante de la matriz $(n-1) \times (n-1)$ que resulta de eliminar la fila i y la columna k de la matriz A. Dado que $(-1)^{i+k}$ es 1 o -1 si $j+k$ es par o impar, podemos desarrollar las siguientes **matrices de signos** para evitar calcular la potencia $(-1)^{i+j}$:

$$\begin{bmatrix} + & - \\ - & + \end{bmatrix}, \quad \begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}, \quad \begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix},$$

etc., que nos indican el signo de la potencia $(-1)^{i+k}$, únicamente viendo la posición i, k para cada matriz del tamaño correspondiente; como fácilmente puede comprobarlo el lector. Entonces para calcular el determinante anteponemos, para cada $k = 1, 2, \ldots, n$, el signo que corresponde al elemento i, k al producto $a_{ik} M_{ik}$ en la matriz de signos del mismo tamaño de A y se suman algebraicamente estos términos desde $k = 1$ hasta $k = n$. Ilustraremos lo anterior en el ejemplo 2.19.

❍ **Nota 2.2** Para formar una matriz de signos se observa lo siguiente:

- Se comienza con el signo $+$ en la primera componente de la primera fila.
- Los signos $+$ y $-$ se van alternando en cada fila y en cada columna.
- Si se han hecho correctamente los dos pasos anteriores, todos los signos en la diagonal serán positivos.

► Ejemplo 2.19

$$\begin{vmatrix} 1 & -1 & 0 & 2 \\ 3 & 2 & -1 & 1 \\ 1 & 2 & 4 & 3 \\ -1 & -2 & 0 & 5 \end{vmatrix} = -(-1)\begin{vmatrix} 1 & -1 & 2 \\ 1 & 2 & 3 \\ -1 & -2 & 5 \end{vmatrix} + (4)\begin{vmatrix} 1 & -1 & 2 \\ 3 & 2 & 1 \\ -1 & -2 & 5 \end{vmatrix}$$

$$= \begin{vmatrix} 2 & 3 \\ -2 & 5 \end{vmatrix} - \begin{vmatrix} -1 & 2 \\ -2 & 5 \end{vmatrix} - \begin{vmatrix} -1 & 2 \\ 2 & 3 \end{vmatrix}$$

$$
\begin{aligned}
&+4\left[\begin{vmatrix} 2 & 1 \\ -2 & 5 \end{vmatrix} + \begin{vmatrix} 3 & 1 \\ -1 & 5 \end{vmatrix} + 2\begin{vmatrix} 3 & 2 \\ -1 & -2 \end{vmatrix}\right] \\
&= 16+1+7+4[12+16-8] \\
&= 104.
\end{aligned}
$$

Donde hemos utilizado la matriz de signos 4×4 y la de 3×3 para saber el signo que debe anteceder al menor cuando se desarrolla el determinante por una fila o columna conveniente.[9] ◀

Calcular el determinante de una matriz por medio de cofactores puede convertirse en un trabajo increíblemente laborioso. Por ejemplo, si una computadora de esta época se utiliza para calcular el determinante de una matriz de tamaño 50×50, nuestro sistema solar habrá desaparecido antes que la computadora dé el resultado final. Por esta razón necesitamos procedimientos que nos auxilien a minimizar los cálculos para encontrar el determinante de una matriz. Nuevamente el método de Gauss será la base para alcanzar ese objetivo y en el siguiente apartado nos abocamos a ese fin.

2.2.2 Propiedades

Enunciamos las siguientes propiedades de los determinantes e invitamos al lector a consultar las correspondientes demostraciones en el apéndice C.

Teorema 2.8 ***(Propiedades de los determinantes)*** *Sean A y B matrices cuadradas de orden n, entonces:*

1. $|AB| = |A||B|$.
2. $|A^t| = |A|$.
3. *Si A es una matriz invertible, entonces* $|A^{-1}| \neq 0$ *y además* $|A^{-1}| = \frac{1}{|A|}$.
4. *Si A es triangular superior o triangular inferior, el determinante de A es el producto de los elementos de la diagonal.*
5. *Supongamos que B es equivalente a la matriz A al aplicarle una operación de renglón:*
   (a) *Si B se obtiene de A mediante la operación de renglón* $R_i \leftrightarrow R_j$, *entonces* $|A| = -|B|$.
   (b) *Si B resulta de A al aplicarle la operación* $R_i \leftrightarrow R_i + \alpha R_j$, *entonces* $|B| = |A|$.
   (c) *Si B se obtiene de A con la operación* $R_i \leftrightarrow \alpha R_i$, $\alpha \neq 0$, *entonces* $|A| = (1/\alpha)|B|$.

❍ **Nota 2.3** De manera análoga a las operaciones de renglón se definen las operaciones de columnas: $G_i \leftrightarrow G_j$, $G_i \leftrightarrow \alpha G_i$ $(\alpha \neq 0)$, $G_i \leftrightarrow \alpha G_i + \beta G_j$ $(\alpha \neq 0)$, que significan, respectivamente, la columna G_i se intercambia con la columna G_j, la columna G_i se cambia por α-veces la columna G_i, la columna G_i se cambia por α-veces la columna G_i más β-veces la columna G_j. Debido a que $|A^t| = |A|$, las propiedades de los determinantes que involucran operaciones en renglones se pueden traducir a propiedades de operaciones en columnas:

[9] Se entiende por fila o columna conveniente, una fila o columna de la matriz que minimice los cálculos; si es que la matriz contiene alguna con esta característica.

1. Si B es la matriz que se obtiene de una matriz A mediante la operación de columna $G_i \leftrightarrow G_j$, entonces $|A| = -|B|$.
2. Si B es la matriz que se obtiene de una matriz A mediante la operación de columna $G_i \leftrightarrow \alpha G_j$, entonces $|A| = (1/\alpha)|B|$.
3. Si B es la matriz que resulta de una matriz A al aplicarle la operación de columna $G_i \leftrightarrow G_i + \beta G_j$, entonces $|A| = |B|$.

▶ **Ejemplo 2.20** Sean $A = \begin{bmatrix} -1 & 3 \\ 4 & 7 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$; entonces $|A| = -19$, $|B| = 5$ y

$$\begin{aligned} |AB| &= \begin{vmatrix} 1 & 8 \\ 15 & 25 \end{vmatrix} \\ &= 25 - 120 \\ &= -95 \\ &= (-19)(5) \\ &= |A|\,|B|. \end{aligned}$$ ◀

▶ **Ejemplo 2.21** Mostrar la propiedad 3.◀

DEMOSTRACIÓN ■ Sea A una matriz de orden n invertible, entonces $AA^{-1} = I_n$. Luego, por la propiedad 4, $|I_n| = 1$ y, por la propiedad 1, $\left|AA^{-1}\right| = |A|\left|A^{-1}\right|$; luego

$$\begin{aligned} 1 &= |I_n| \\ &= \left|AA^{-1}\right| \\ &= |A|\left|A^{-1}\right|; \end{aligned}$$

de donde $|A| \neq 0$ y $\left|A^{-1}\right| = 1/|A|$. ■

▶ **Ejemplo 2.22** Por la propiedad 4

$$\begin{vmatrix} -1 & 3 & 4 & -1 \\ 0 & 2 & -2 & 3 \\ 0 & 0 & 3 & -7 \\ 0 & 0 & 0 & 4 \end{vmatrix} = (-1)(2)(3)(4) = -24.$$ ◀

▶ **Ejemplo 2.23** Por las propiedades 5*(a)* y 4

$$\begin{vmatrix} -1 & 3 & 2 \\ 0 & 0 & 1 \\ 0 & 3 & -5 \end{vmatrix} = -\begin{vmatrix} -1 & 3 & 2 \\ 0 & 3 & -5 \\ 0 & 0 & 1 \end{vmatrix} = -(-3) = 3.$$

Comprobación:

$$\begin{vmatrix} -1 & 3 & 2 \\ 0 & 0 & 1 \\ 0 & 3 & -5 \end{vmatrix} = -(1)\begin{vmatrix} -1 & 3 \\ 0 & 3 \end{vmatrix} = -3.$$ ◀

▶ **Ejemplo 2.24** Por las propiedades 5*(b)* y 4

$$\begin{vmatrix} 1 & -2 & 1 \\ -5 & 5 & -1 \\ 0 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 1 & -2 & 1 \\ 0 & -5 & 4 \\ 0 & 0 & 3 \end{vmatrix} = -15. ◀$$

▶ **Ejemplo 2.25** Por la propiedad 5*(c)*

$$\begin{vmatrix} 4 & -4 & 8 \\ 1 & -1 & 2 \\ 1 & 1 & 3 \end{vmatrix} = \frac{1}{1/4}\begin{vmatrix} 1 & -2 & 2 \\ 1 & -1 & 2 \\ 1 & 1 & 3 \end{vmatrix} = 4\begin{vmatrix} 1 & -2 & 2 \\ 1 & -1 & 2 \\ 1 & 1 & 3 \end{vmatrix}$$

y por la propiedad 5*(b)*

$$\begin{aligned} \begin{vmatrix} 4 & -8 & 8 \\ 1 & -1 & 2 \\ 1 & 1 & 3 \end{vmatrix} &= 4\begin{vmatrix} 1 & -2 & 2 \\ 1 & -1 & 2 \\ 1 & 1 & 3 \end{vmatrix} \\ &= 4\begin{vmatrix} 1 & -2 & 2 \\ 0 & 1 & 0 \\ 1 & 1 & 3 \end{vmatrix} \\ &= 4(1)\begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} \\ &= 4. ◀ \end{aligned}$$

▶ **Ejemplo 2.26** Calcular el determinante de la matriz

$$\begin{bmatrix} 3 & -6 & 3 & 9 \\ 2 & 1 & -2 & 1 \\ 3 & 2 & -4 & 2 \\ 5 & 1 & 2 & 3 \end{bmatrix}$$

llevándola a la forma triangular superior mediante el método de Gauss y haciendo uso de las propiedades 5*(a)*, *(b)*, *(c)* y 4. ◀

Solución

$$\begin{aligned} \begin{vmatrix} 3 & -6 & 3 & 9 \\ 2 & 1 & -2 & 1 \\ 3 & 2 & -4 & 2 \\ 5 & 1 & 2 & 3 \end{vmatrix} &= 3\begin{vmatrix} 1 & -2 & 1 & 3 \\ 2 & -1 & -2 & 1 \\ 3 & 2 & -4 & 2 \\ 5 & 1 & 2 & 3 \end{vmatrix} \\ &= 3\begin{vmatrix} 1 & -2 & 1 & 3 \\ 0 & 5 & -4 & -5 \\ 0 & -8 & -7 & -7 \\ 0 & 11 & -3 & -12 \end{vmatrix} \\ &= 3\begin{vmatrix} 1 & -2 & 1 & 3 \\ 0 & 5 & -4 & -5 \\ 0 & 8 & -7 & -7 \\ 0 & 1 & 5 & -2 \end{vmatrix} \\ &= -3\begin{vmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & 5 & -2 \\ 0 & 8 & -7 & -7 \\ 0 & 5 & -4 & -5 \end{vmatrix} \end{aligned}$$

$$= -3\begin{vmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & 5 & -2 \\ 0 & 0 & -47 & 9 \\ 0 & 0 & -29 & 5 \end{vmatrix}$$

$$= (-3)(-47)\begin{vmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & 5 & -2 \\ 0 & 0 & 1 & -9/47 \\ 0 & 0 & -29 & 5 \end{vmatrix}$$

$$= (-3)(-47)\begin{vmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & 5 & -2 \\ 0 & 0 & 1 & -9/47 \\ 0 & 0 & 0 & -26/47 \end{vmatrix}$$

$$= (-3)(-47)(-26/47)$$

$$= -78. \quad ✔$$

2.2.3 Método de la adjunta para hallar la inversa

A continuación daremos un método alternativo para hallar la inversa de una matriz. Debemos aclarar que, en comparación con el método de Gauss-Jordan, resulta ser muy ineficiente; pues para llevarlo a efecto se tienen que calcular varios determinantes. Sin embargo, las implicaciones teóricas y prácticas de dicho método justifican muy bien su estudio.

Definición 2.6 *Sea A una matriz cuadrada de orden n. Se definen:*

- *La* **matriz de cofactores** *de A como $\operatorname{Cof}(A) = [c_{ij}]$, donde c_{ij} es el cofactor del elemento ij de A.*
- *La matriz* **adjunta** *de A por $\operatorname{Adj}(A) = [\operatorname{Cof}(A)]^t$; es decir, la matriz adjunta de A es la transpuesta de la matriz de cofactores.*

Vimos en el inciso 3 de las propiedades de los determinates (cfr. teorema 2.8, pág. 80), que si una matriz es invertible su determinante es diferente de cero; el recíproco es cierto y está contenido en el teorema 2.9; y se invita al lector a consultar su demostración en el apéndice C.

Teorema 2.9 ***(Método de la adjunta)*** *Si $A \in \mathfrak{M}_{n\times n}$, entonces*

$$A\operatorname{Adj}(A) = \det(A)I_n.$$

Luego, A es invertible si y sólo si $\det(A) \neq 0$ y en tal caso

$$A^{-1} = \frac{1}{|A|}\operatorname{Adj}(A).$$

► **Ejemplo 2.27** Sea $A = \begin{bmatrix} -1 & 1 & 2 \\ 2 & -1 & 1 \\ -2 & -3 & 4 \end{bmatrix}$, entonces $c_{11} = -1$, $c_{12} = -10$, $c_{13} = -8$, $c_{21} = -10$, $c_{22} = 0$, $c_{23} = -5$, $c_{31} = 3$, $c_{32} = 5$, $c_{33} = -1$; por tanto,

$$\operatorname{Cof}(A) = \begin{bmatrix} -1 & -10 & -8 \\ -10 & 0 & -5 \\ 3 & 5 & -1 \end{bmatrix},$$

$$\operatorname{Adj}(A) = \begin{bmatrix} -1 & -10 & 3 \\ -10 & 0 & 5 \\ -8 & -5 & -1 \end{bmatrix},$$

y $|A| = -25$. Luego,

$$A^{-1} = \frac{1}{|A|}\operatorname{Adj}(A) = -\frac{1}{25}\begin{bmatrix} -1 & -10 & 3 \\ -10 & 0 & -5 \\ -8 & -5 & -1 \end{bmatrix} = \begin{bmatrix} 1/25 & 2/5 & -3/25 \\ 2/5 & 0 & -1/5 \\ 8/25 & 1/5 & 1/25 \end{bmatrix}.$$ ◄

Con el teorema 2.9 es posible añadir una condición más de equivalencia a los teoremas 2.3 y 2.6 (cfr. págs. 67 y 73) que involucra ahora el determinante.

Teorema 2.10 *Si $A \in \mathfrak{M}_{n\times n}$ las siguientes condiciones son equivalentes a pares:*

1. *El sistema $A\vec{x} = \vec{b}$ tiene solución única $\forall \vec{b}$.*
2. *$A \sim I_n$.*
3. *A es invertible.*
4. *A es producto de matrices elementales.*
5. $\det(A) \neq 0$.

2.2.4 Regla de Cramer

En este breve apartado explicamos un algoritmo que al lector le será familiar de sus cursos de bachillerato; el cual da un método para hallar la solución de un sistema cuadrado utilizando determinantes. Nuevamente cabe señalar que su utilidad práctica directa es muy raquítica para sistemas con más de cuatro variables, pero su conocimiento es de gran interés teórico e histórico y tiene importantes implicaciones indirectas de valor práctico.[10]

Teorema 2.11 *Sea A una matriz cuadrada de orden n. Sea, para cada $i = 1,2,\ldots,n$, Δ_i el determinante que resulta de la matriz que se obtiene de A al sustituir la i-ésima columna por $\vec{b}$. Entonces, si $\Delta = |A| \neq 0$ el sistema $A\vec{x} = \vec{b}$ tiene solución única*

$$x_i = \frac{\Delta_i}{\Delta},$$

para $i = 1,2,\ldots,n$.

[10] Para una demostración del teorema 2.11, consulte el apéndice C.

▶ **Ejemplo 2.28** Resolver el sistema

$$\begin{aligned} 3x_1 - 5x_2 &= -2 \\ -2x_1 + x_2 &= -3. \end{aligned}$$

En este caso, $A = \begin{bmatrix} 3 & -5 \\ -2 & 1 \end{bmatrix}$; $\vec{b} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$.

$$\begin{aligned} \Delta &= \begin{vmatrix} 3 & -5 \\ -2 & 1 \end{vmatrix} = -7, \\ \Delta_1 &= \begin{vmatrix} -2 & -5 \\ -3 & 1 \end{vmatrix} = -17, \\ \Delta_2 &= \begin{vmatrix} 3 & -2 \\ -2 & -3 \end{vmatrix} = -13. \end{aligned}$$

Luego,

$$\begin{aligned} x_1 &= \frac{\Delta_1}{\Delta} = \frac{17}{7}, \\ x_2 &= \frac{\Delta_2}{\Delta} = \frac{13}{7}. \end{aligned}$$ ◀

❍ **Nota 2.4** Si $\Delta = 0$, un sistema cuadrado $A\vec{x} = \vec{b}$ puede o no tener solución. Por ejemplo, en los sistemas

$$\begin{aligned} x_1 + x_2 &= 1 \\ 2x_1 + 2x_2 &= 2 \end{aligned} \quad \text{y} \quad \begin{aligned} x_1 + x_2 &= 1 \\ 2x_1 + 2x_2 &= 0 \end{aligned}$$

$\Delta = 0$, pero el primero tiene una infinidad de soluciones y el segundo es inconsistente.

2.2.5 Determinantes de matrices con componentes complejas

Para calcular determinantes de matrices con componentes complejas, como hemos hecho antes en estos casos, aplicaremos todas las técnicas que estudiamos en esta sección para el caso de determinantes de matrices reales pero utilizando el álgebra de los números complejos.[11] Otra vez toda la teoría que desarrollamos para determinantes de matrices reales es también válida para el caso de matrices en $\mathfrak{M}_n(\mathbb{C})$.

▶ **Ejemplo 2.29** Calcular el determinante de la matriz

$$A = \begin{bmatrix} 1-i & 0 & 2i & 0 \\ 3+2i & -1-4i & 0 & 1 \\ -i & 2 & 3-3i & -5i \\ 0 & -1 & 3 & 4 \end{bmatrix}.$$ ◀

[11] Cfr. B1 del apéndice B, 1.1.5, 1.2.8, 2.1.5.

Solución Desarrollemos el determinante por la primera fila mediante cofactores.

$$\begin{aligned}
|A| &= (1-i)\begin{vmatrix} -1-4i & 0 & 1 \\ 2 & 3-3i & -5i \\ -1 & 3 & 4 \end{vmatrix} + (2i)\begin{vmatrix} 3+2i & -1-4i & 1 \\ -i & 2 & -5i \\ 0 & -1 & 4 \end{vmatrix} \\
&= (1-i)\left[(-1-4i)\begin{vmatrix} 3-3i & -5i \\ 3 & 4 \end{vmatrix} + (1)\begin{vmatrix} 2 & 3-3i \\ -1 & 3 \end{vmatrix}\right] \\
&\quad +2i\left[-(-1)\begin{vmatrix} 3+2i & 1 \\ -i & -5i \end{vmatrix} + (4)\begin{vmatrix} 3+2i & -1-4i \\ -i & 2 \end{vmatrix}\right] \\
&= (1-i)\,[(-1-4i)(12-12i+15i)+6+3-3i] \\
&\quad +2i\,[10-14i+4(6+4i+4-i)] \\
&= (1-i)\,[(-1-4i)(12+3i)+9-3i]+2i\,[10-14i+40+12i] \\
&= (1-i)(9-54i)+2i(50-2i) \\
&= -45-63i+4+100i \\
&= -41+37i. \quad ✔
\end{aligned}$$

▶ **Ejemplo 2.30** Calcular la inversa de la matriz $A = \begin{bmatrix} 2+i & 3i \\ -i & 1+4i \end{bmatrix}$ por medio de la adjunta.◀

Solución

$$\begin{aligned}
\det(A) &= (2+i)(1+4i)-(3i)(-i) \\
&= -2+9i-3 \\
&= -5+9i
\end{aligned}$$

luego,

$$\begin{aligned}
A^{-1} &= \frac{1}{-5+9i}\begin{bmatrix} 1+4i & -3i \\ i & 2+i \end{bmatrix} \\
&= \frac{1}{106}\begin{bmatrix} 31-29i & -27+15i \\ 9-5i & -1-23i \end{bmatrix}. \quad ✔
\end{aligned}$$

2.3 Ejercicios resueltos y ejercicios propuestos

2.3.1 Ejercicios resueltos

Matrices invertibles y sus inversas

1 Comprobar si la matriz C es la matriz inversa de la matriz A en cada caso.

(a) $A = \begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix}, C = \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}.$

(b) $A = \begin{bmatrix} -1 & 2 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix}, C = \begin{bmatrix} -2 & 3 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}.$

(c) $A=\begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & -1 & 3 & 2 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 1 & 1 \end{bmatrix}, C=\begin{bmatrix} -1 & 1 & 0 & -1 \\ -2 & 1 & 1 & -3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & -1 & 4 \end{bmatrix}.$

(d) $A=\begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}, C=\begin{bmatrix} -1 & -3 \\ 1 & -2 \end{bmatrix}.$

Solución Si $AC=I$, se tiene $C=A^{-1}$.

(a)
$$\begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Por tanto, $C=A^{-1}$.

(b)
$$\begin{bmatrix} -1 & 2 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix}\begin{bmatrix} -2 & 3 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Por tanto, $C=A^{-1}$.

(c)
$$\begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & -1 & 3 & 2 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 & 0 & -1 \\ -2 & 1 & 1 & -3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & -1 & 4 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Por tanto, $C=A^{-1}$.

(d)
$$\begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & -3 \\ 1 & -2 \end{bmatrix}=\begin{bmatrix} 1 & -12 \\ 0 & -5 \end{bmatrix}\neq\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Por tanto, $C\neq A^{-1}$. ✔

2 Utilizar las matrices A y C del ejercicio 1 para resolver los sistemas lineales.

(a)
$$\begin{aligned} -x+2y&=-1 \\ -2x+3y&=3 \end{aligned}$$

(b)
$$\begin{aligned} -x+2y-z &= -2 \\ y &= 4 \\ x-y+2z &= -3 \end{aligned}$$

(c)
$$\begin{aligned} x_1-x_2+2x_3+x_4 &= -1 \\ 2x_1-x_2+3x_3+2x_4 &= 4 \\ 4x_3+3x_4 &= -6 \\ x_3+x_4 &= 2 \end{aligned}$$

Solución Si una matriz cuadrada es invertible, entonces el sistema $A\vec{x}=\vec{b}$ tiene solución única $\vec{x}=A^{-1}\vec{b}$.

(a) La matriz de coeficientes para este sistema es
$$A=\begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix}$$

y, por el primer inciso del ejercicio 1, $A^{-1} = \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}$; por tanto,

$$\vec{x} = \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} -9 \\ -5 \end{bmatrix}.$$

(b) Para este sistema la matriz de coeficientes es $A = \begin{bmatrix} -1 & 2 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix}$ y, por el segundo inciso del ejercicio 1, $A^{-1} = \begin{bmatrix} -2 & 3 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}$; así que,

$$\vec{x} = \begin{bmatrix} -2 & 3 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 4 \\ -3 \end{bmatrix} = \begin{bmatrix} 19 \\ 4 \\ -9 \end{bmatrix}.$$

(c) En este caso la matriz de coeficientes es

$$A = \begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & -1 & 3 & 2 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

y por el tercer inciso del ejercicio 1

$$\vec{x} = \begin{bmatrix} -1 & 1 & 0 & -1 \\ -2 & 1 & 1 & -3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 4 \\ -6 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ -6 \\ -12 \\ 14 \end{bmatrix}.$$ ✔

En los ejercicios 3 y 4, determinar si la matriz dada es invertible.

3 $A = \begin{bmatrix} 1 & -2 & 3 \\ 3 & -4 & 5 \\ -2 & 1 & 6 \end{bmatrix}$.

Solución Una matriz es invertible si es equivalente (por filas) a la identidad (cfr. teorema 2.3)

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 3 & -4 & 5 \\ -2 & 1 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & -4 \\ 0 & -3 & 12 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & -4 \\ 0 & 0 & 12 \end{bmatrix}.$$

Puesto que todas las columnas en la última matriz en forma escalonada equivalente a A tienen pivote, se deduce que A es equivalente a la identidad y, por tanto, es invertible (cfr. nota 1.6). ✔

4 $B = \begin{bmatrix} 1 & -2 & 3 & -1 \\ -2 & 1 & 2 & 1 \\ 3 & -3 & 2 & 4 \\ 1 & 1 & -4 & 6 \end{bmatrix}$

Solución Una matriz es invertible si es equivalente a la identidad.

$$B = \begin{bmatrix} 1 & -2 & 3 & -1 \\ -2 & 1 & 2 & 1 \\ 3 & -3 & 2 & 4 \\ 1 & 1 & -4 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 & -1 \\ 0 & -3 & 8 & -1 \\ 0 & 3 & -7 & 7 \\ 0 & 3 & -7 & 7 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -2 & 3 & -1 \\ 0 & -3 & 8 & -1 \\ 0 & 3 & -7 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 & -1 \\ 0 & -3 & 8 & -1 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix} = F.$$

Dado que la última columna de la matriz en forma escalonada F equivalente a B no tiene pivote, se deduce que B no puede ser equivalente a la identidad. Por tanto, B no es invertible. ✔

5 Encontrar los valores de r para que la matriz $A = \begin{bmatrix} 1 & -2 & -2 \\ -2 & r & 1 \\ 3 & -1 & r \end{bmatrix}$ sea invertible.

Solución Para que la matriz A sea invertible es necesario y suficiente que sea equivalente a la identidad. Ya que

$$A = \begin{bmatrix} 1 & -2 & -2 \\ -2 & r & 1 \\ 3 & -5 & r \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -2 \\ 0 & r-4 & -3 \\ 0 & 1 & 6+r \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -2 \\ 0 & 1 & 6+r \\ 0 & r-4 & -3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -2 & -2 \\ 0 & 1 & 6+r \\ 0 & 0 & 21-2r-r^2 \end{bmatrix}$$

y la última matriz (que está en forma escalonada equivalente a A) es equivalente a la identidad si $21 - 2r - r^2 \neq 0$, A es invertible si $r \neq -1 \pm \sqrt{22}$. ✔

En los ejercicios 6 y 7 calcular la inversa, si existe, de cada matriz por el método de Gauss-Jordan.

6 $\begin{bmatrix} -1 & 2 & 1 \\ 3 & 2 & -1 \\ 5 & -1 & 2 \end{bmatrix}$

Solución

$$\left[\begin{array}{rrr|rrr} -1 & 2 & 1 & 1 & 0 & 0 \\ 3 & 2 & -1 & 0 & 1 & 0 \\ 5 & -1 & 2 & 0 & 0 & 1 \end{array}\right] \xleftrightarrow[R_3 \leftrightarrow 5R_1 + R_3]{R_2 \leftrightarrow 3R_1 + R_2} \left[\begin{array}{rrr|rrr} -1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 8 & 2 & 3 & 1 & 0 \\ 0 & 9 & 7 & 5 & 0 & 1 \end{array}\right]$$

$$\xleftrightarrow[R_2 \leftrightarrow R_3 - R_2]{} \left[\begin{array}{rrr|rrr} -1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 5 & 2 & -1 & 1 \\ 0 & 9 & 7 & 5 & 0 & 1 \end{array}\right]$$

$$\xleftrightarrow[R_3 \leftrightarrow -9R_2 + R_3]{} \left[\begin{array}{rrr|rrr} -1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 5 & 2 & -1 & 1 \\ 0 & 0 & -38 & -13 & 9 & -8 \end{array}\right]$$

$$\xleftrightarrow[R_3 \leftrightarrow -\frac{1}{38}R_3]{R_1 \leftrightarrow -R_1} \left[\begin{array}{rrr|rrr} 1 & -2 & -1 & -1 & 0 & 0 \\ 0 & 1 & 5 & 2 & -1 & 1 \\ 0 & 0 & 1 & \frac{13}{38} & -\frac{9}{38} & \frac{4}{19} \end{array}\right]$$

$$\xleftrightarrow[R_1 \leftrightarrow R_3+R_1]{R_2 \leftrightarrow -5R_3+R_2} \left[\begin{array}{ccc|ccc} 1 & -2 & 0 & -\frac{25}{38} & -\frac{9}{38} & \frac{4}{19} \\ 0 & 1 & 0 & \frac{11}{38} & \frac{7}{38} & -\frac{1}{19} \\ 0 & 0 & 1 & \frac{13}{38} & -\frac{9}{38} & \frac{4}{19} \end{array}\right]$$

$$\xleftrightarrow{R_1 \leftrightarrow 2R_2+R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{3}{38} & \frac{5}{38} & \frac{2}{19} \\ 0 & 1 & 0 & \frac{11}{38} & \frac{7}{38} & -\frac{1}{19} \\ 0 & 0 & 1 & \frac{13}{38} & -\frac{9}{38} & \frac{4}{19} \end{array}\right]$$

Entonces,

$$\begin{bmatrix} -1 & 2 & 1 \\ 3 & 2 & -1 \\ 5 & -1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{3}{38} & \frac{5}{38} & \frac{2}{19} \\ \frac{11}{38} & \frac{7}{38} & -\frac{1}{19} \\ \frac{13}{38} & -\frac{9}{38} & \frac{4}{19} \end{bmatrix}.$$

Comprobación:

$$\begin{bmatrix} -1 & 2 & 1 \\ 3 & 2 & -1 \\ 5 & -1 & 2 \end{bmatrix} \begin{bmatrix} -\frac{3}{38} & \frac{5}{38} & \frac{2}{19} \\ \frac{11}{38} & \frac{7}{38} & -\frac{1}{19} \\ \frac{13}{38} & -\frac{9}{38} & \frac{4}{19} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \checkmark$$

7 $\begin{bmatrix} 2 & 1 & 1 & -1 \\ 3 & -1 & 1 & 2 \\ 5 & 2 & -1 & 3 \\ -7 & 2 & 1 & 2 \end{bmatrix}$

Solución

$$\left[\begin{array}{cccc|cccc} 2 & 1 & 1 & -1 & 1 & 0 & 0 & 0 \\ 3 & -1 & 1 & 2 & 0 & 1 & 0 & 0 \\ 5 & 2 & -1 & 3 & 0 & 0 & 1 & 0 \\ -7 & 2 & 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cccc|cccc} 1 & -2 & 0 & 3 & -1 & 1 & 0 & 0 \\ 3 & -1 & 1 & 2 & 0 & 1 & 0 & 0 \\ 5 & 2 & -1 & 3 & 0 & 0 & 1 & 0 \\ -7 & 2 & 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|cccc} 1 & -2 & 0 & 3 & -1 & 1 & 0 & 0 \\ 0 & 5 & 1 & -7 & 3 & -2 & 0 & 0 \\ 0 & 12 & -1 & -12 & 5 & -5 & 1 & 0 \\ 0 & -12 & 1 & 23 & -7 & 7 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|cccc} 1 & -2 & 0 & 3 & -1 & 1 & 0 & 0 \\ 0 & 5 & 1 & -7 & 3 & -2 & 0 & 0 \\ 0 & 12 & -1 & -12 & 5 & -5 & 1 & 0 \\ 0 & 0 & 0 & 11 & -2 & 2 & 1 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|cccc} 1 & -2 & 0 & 3 & -1 & 1 & 0 & 0 \\ 0 & 5 & 1 & -7 & 3 & -2 & 0 & 0 \\ 0 & 0 & -17 & 24 & -11 & -1 & 5 & 0 \\ 0 & 0 & 0 & 11 & -2 & 2 & 1 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|cccc} 1 & -2 & 0 & 3 & -1 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{5} & -\frac{7}{5} & \frac{3}{5} & -\frac{2}{5} & 0 & 0 \\ 0 & 0 & 1 & -\frac{24}{17} & \frac{11}{17} & \frac{1}{17} & -\frac{5}{17} & 0 \\ 0 & 0 & 0 & 1 & -\frac{2}{11} & \frac{2}{11} & \frac{1}{11} & \frac{1}{11} \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|cccc} 1 & -2 & 0 & 0 & -\frac{5}{11} & \frac{5}{11} & -\frac{3}{11} & -\frac{3}{11} \\ 0 & 1 & \frac{1}{5} & 0 & \frac{19}{55} & -\frac{8}{55} & \frac{7}{55} & \frac{7}{55} \\ 0 & 0 & 1 & 0 & \frac{73}{187} & \frac{59}{187} & -\frac{31}{187} & \frac{24}{187} \\ 0 & 0 & 0 & 1 & -\frac{2}{11} & \frac{2}{11} & \frac{1}{11} & \frac{1}{11} \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & \frac{15}{187} & \frac{7}{187} & \frac{9}{187} & -\frac{13}{187} \\ 0 & 1 & 0 & 0 & \frac{50}{187} & -\frac{39}{187} & \frac{30}{187} & \frac{19}{187} \\ 0 & 0 & 1 & 0 & \frac{73}{187} & \frac{59}{187} & -\frac{31}{187} & \frac{24}{187} \\ 0 & 0 & 0 & 1 & -\frac{2}{11} & \frac{2}{11} & \frac{1}{11} & \frac{1}{11} \end{array}\right].$$

Por tanto,

$$\begin{bmatrix} 2 & 1 & 1 & -1 \\ 3 & -1 & 1 & 2 \\ 5 & 2 & -1 & 3 \\ -7 & 2 & 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{15}{187} & \frac{7}{187} & \frac{9}{187} & -\frac{13}{187} \\ \frac{50}{187} & -\frac{39}{187} & \frac{30}{187} & \frac{19}{187} \\ \frac{73}{187} & \frac{59}{187} & -\frac{31}{187} & \frac{24}{187} \\ -\frac{2}{11} & \frac{2}{11} & \frac{1}{11} & \frac{1}{11} \end{bmatrix}.$$

Comprobación:

$$\begin{bmatrix} 2 & 1 & 1 & -1 \\ 3 & -1 & 1 & 2 \\ 5 & 2 & -1 & 3 \\ -7 & 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} \frac{15}{187} & \frac{7}{187} & \frac{9}{187} & -\frac{13}{187} \\ \frac{50}{187} & -\frac{39}{187} & \frac{30}{187} & \frac{19}{187} \\ \frac{73}{187} & \frac{59}{187} & -\frac{31}{187} & \frac{24}{187} \\ -\frac{2}{11} & \frac{2}{11} & \frac{1}{11} & \frac{1}{11} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$ ✔

En los ejercicios 8 a 10, calcular la inversa de cada matriz elemental utilizando la explicación dada antes del teorema 2.5 (cfr. pág. 72).

8 $E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$

Solución La matriz E se obtuvo de la identidad I_4 al aplicarle la operación de renglón $R_3 \leftrightarrow R_2$. Entonces la matriz inversa de E se obtiene aplicando a I_4 la operación de renglón $R_2 \leftrightarrow R_3$; por tanto, $E^{-1} = E$. ✔

9 $E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Solución La matriz E se obtuvo de I_4 aplicándole la operación de renglón $R_2 \leftrightarrow -3R_2$; por tanto, la matriz inversa de E se calcula aplicando la operación de renglón $R_2 \leftrightarrow (-1/3)R_2$ a I_4; esto es,

$$E^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$ ✔

10 $E=\begin{bmatrix}1&0&0&0\\0&-2&3&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.$

Solución La matriz elemental E es el resultado de aplicar la operación de renglón $R_2\leftrightarrow -2R_2+3R_3$ a la matriz I_4; por ende, la matriz inversa de E se obtiene aplicando la operación $R_2\leftrightarrow(-1/2)R_2+(3/2)R_3$ a I_4; o sea

$$E^{-1}=\begin{bmatrix}1&0&0&0\\0&-\frac{1}{2}&\frac{3}{2}&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.\ ✔$$

11 Encontrar la inversa de la matriz $A=\begin{bmatrix}1&-1&1\\2&1&1\\2&-1&2\end{bmatrix}$ y escribir A^{-1} como producto de matrices elementales.

Solución
$$\left[\begin{array}{rrr|rrr}1&-1&1&1&0&0\\2&1&10&1&0&\\2&-1&2&0&0&1\end{array}\right]\xleftrightarrow[R_3\leftrightarrow-2R_1+R_3]{R_2\leftrightarrow-2R_1+R_2}\left[\begin{array}{rrr|rrr}1&-1&1&1&0&0\\0&3&-1&-2&1&0\\0&1&0&-2&0&1\end{array}\right]$$

$$\xleftrightarrow[R_3\leftrightarrow R_2]{}\left[\begin{array}{rrr|rrr}1&-1&1&1&0&0\\0&1&0&-2&0&1\\0&3&-1&-2&1&0\end{array}\right]$$

$$\xleftrightarrow[R_3\leftrightarrow-3R_2+R_3]{}\left[\begin{array}{rrr|rrr}1&-1&1&1&0&0\\0&1&0&-2&0&1\\0&0&-1&4&1&-3\end{array}\right]$$

$$\xleftrightarrow[R_1\leftrightarrow R_3+R_1]{}\left[\begin{array}{rrr|rrr}1&-1&0&5&1&-3\\0&1&0&-2&0&1\\0&0&-1&4&1&-3\end{array}\right]$$

$$\xleftrightarrow[R_1\leftrightarrow R_2+R_1]{}\left[\begin{array}{rrr|rrr}1&0&0&3&1&-2\\0&1&0&-2&0&1\\0&0&-1&4&1&-3\end{array}\right]$$

$$\xleftrightarrow[R_3\leftrightarrow-R_3]{}\left[\begin{array}{rrr|rrr}1&0&0&3&1&-2\\0&1&0&-2&0&1\\0&0&1&-4&-1&3\end{array}\right]$$

Con lo que
$$A^{-1}=\begin{bmatrix}3&1&-2\\-2&0&1\\-4&-1&3\end{bmatrix}.$$

Las matrices elementales que corresponden a cada operación de renglón (cfr. teorema 2.4) son las que se obtienen al aplicar las mismas operaciones de fila a I_3 que las que se aplican en forma sucesiva a la matriz A para obtener A^{-1}: $E_1=\begin{bmatrix}1&0&0\\-2&1&0\\0&0&1\end{bmatrix}$ $(R_2\leftrightarrow-2R_1+R_2)$; $E_2=\begin{bmatrix}1&0&0\\0&1&0\\-2&0&1\end{bmatrix}$ $(R_3\leftrightarrow-2R_1+R_3)$;

$$E_3=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}(R_3\leftrightarrow R_2);\ E_4=\begin{bmatrix}1&0&0\\0&1&0\\0&-3&1\end{bmatrix}(R_3\leftrightarrow -3R_2+R_3);\ E_5=\begin{bmatrix}1&0&1\\0&1&0\\0&0&1\end{bmatrix}$$

$$(R_1\leftrightarrow R_3+R_1);\ E_6=\begin{bmatrix}1&1&0\\0&1&0\\0&0&1\end{bmatrix}(R_1\leftrightarrow R_2+R_1);\ E_7=\begin{bmatrix}1&0&0\\0&1&0\\0&0&-1\end{bmatrix}(R_3\leftrightarrow -R_3).$$

Entonces, $E_7E_6E_5E_4E_3E_2E_1=A^{-1}$ como fácilmente se puede comprobar realizando el producto de las matrices elementales. ✔

12 Sea $A=\begin{bmatrix}1&-2&1\\-1&1&2\\1&-2&3\end{bmatrix}$. Encontrar, si es que existe, una matriz B tal que $AB=A^2+3A$.

Solución Si la matriz A es invertible, entonces

$$\begin{aligned}AB &= A^2+3A\Rightarrow A^{-1}(AB)=A^{-1}(A^2+3A)=(A^{-1}A)A+3A^{-1}A\\&\Rightarrow I_3B=I_3A+3I_3\Rightarrow B=A+3I_3.\end{aligned}$$

Luego, el problema tiene solución si A es invertible; y puesto que

$$\begin{bmatrix}1&-2&1\\-1&1&2\\1&-2&3\end{bmatrix}\sim\begin{bmatrix}1&-2&1\\0&-1&3\\0&0&2\end{bmatrix},$$

A es equivalente a la identidad y, por tanto, es invertible. Así que la solución existe y está dada por

$$B=A+3I_3=\begin{bmatrix}4&-2&1\\-1&4&2\\1&-2&6\end{bmatrix}.\quad ✔$$

13 Resolver el ejercicio 12 con la matriz $A=\begin{bmatrix}0&1\\0&0\end{bmatrix}$.

Solución En este caso no se puede proceder como en el ejercicio 12, pues la matriz A no es invertible (está en forma escalonada y no tiene pivote en la primera columna). Sin embargo,

$$A^2+3A=\begin{bmatrix}0&3\\0&0\end{bmatrix}$$

por lo que es fácil encontrar solución a la ecuación $A^2+3A=AB$ por inspección: $B=\begin{bmatrix}0&0\\0&3\end{bmatrix}$. En efecto,

$$AB=\begin{bmatrix}0&3\\0&0\end{bmatrix}.\quad ✔$$

14 Determinar si la siguiente matriz es invertible y de ser así calcular su inversa.

$$A=\begin{bmatrix} 1 & -i & 1+i & -2 \\ i & 2 & 2i & 4 \\ 1+i & i & -1 & 1 \\ -2 & i & 3 & 4 \end{bmatrix}.$$

Solución Se aplica el método de Gauss-Jordan:

$$\left[\begin{array}{cccc|cccc} 1 & -i & 1+i & -2 & 1 & 0 & 0 & 0 \\ i & 2 & 2i & 4 & 0 & 1 & 0 & 0 \\ 1+i & i & -1 & 1 & 0 & 0 & 1 & 0 \\ -2 & i & 3 & 4 & 0 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cccc|cccc} 1 & -i & 1+i & -21 & 0 & 0 & 0 & \\ 0 & i & -1+i & -2+4i1 & i & 0 & 0 & \\ 0 & -1+2i & -1-2i & 3+2i-1-i & 0 & 1 & 0 & \\ 0 & -i & -i & 0 & 2 & 0 & 0 & 1 \end{array}\right] \sim$$

$$\left[\begin{array}{cccc|cccc} 1 & -i & 1+i & -2 & 1 & 0 & 0 & 0 \\ 0 & i & -1+i & -2+4i & 1 & i & 0 & 0 \\ 0 & 0 & 3+2i & 4+11i & 2-3i & 2+i & i & 0 \\ 0 & 0 & 4+3i & -2+4i & 3 & i & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cccc|cccc} 1 & -i & 1+i & -2 & 1 & 0 & 0 & 0 \\ 0 & i & -1+i & -2+4i & 1 & i & 0 & 0 \\ 0 & 0 & 3+2i & 4+11i & 2-3i & 2+i & i & 0 \\ 0 & 0 & 0 & -3+48i & 8-12i & 7+7i & -3+4i & -3-2i \end{array}\right] \sim$$

$$\left[\begin{array}{cccc|cccc} 3-48i & -48-3i & 51-45i & 0 & -13-24i & -14+14i & 6-8i & 6+4i \\ 0 & 48+3i & 45+51i & 0 & 35+8i & 6+17i & -10-20i & 14-8i \\ 0 & 0 & 105-138i & 0 & 26-65i & 5+12i & -8-14i & 10-41i \\ 0 & 0 & 0 & -3+48i & 8-12i & 7+7i & -3+4i & -3-2i \end{array}\right] \sim$$

$$\left[\begin{array}{cccc|cccc} 6309+5454i & 5454-6309i & 0 & 0 & 3078-3759i & 4197-75i & -564+1314i & -2517-2133i \\ 0 & -5454+6309i & 0 & 0 & -294+2391i & -3363-162i & 4164-318i & 2175+1437i \\ 0 & 0 & 105-138i & 0 & 26-65i & 5+12i & -8-14i & 10-41i \\ 0 & 0 & 0 & -3+48i & 8-12i & 7+7i & -3+4i & -3-2i \end{array}\right] \sim$$

$$\left[\begin{array}{cccc|cccc} 6309+5454i & 0 & 0 & 0 & 2784-1368i & 834-237i & 3600+996i & -342-696i \\ 0 & -5454+6309i & 0 & 0 & -294+2391i & -3363-162i & 4164-318i & 2175+1437i \\ 0 & 0 & 105-138i & 0 & 26-65i & 5+12i & -8-14i & 10-41i \\ 0 & 0 & 0 & -3+48i & 8-12i & 7+7i & -3+4i & -3-2i \end{array}\right] \sim$$

$$\left[\begin{array}{cccc|cccc} 6309+5454i & 0 & 0 & 0 & 2784-1368i & 834-237i & 3600+996i & -342-696i \\ 0 & -5454+6309i & 0 & 0 & -294+2391i & -3363-162i & 4164-318i & 2175+1437i \\ 0 & 0 & 105-138i & 0 & 26-65i & 5+12i & -8-14i & 10-41i \\ 0 & 0 & 0 & -3+48i & 8-12i & 7+7i & -3+4i & -3-2i \end{array}\right] \sim$$

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & \frac{112}{771}-\frac{88}{257}i & \frac{44}{771}-\frac{67}{771}i & \frac{104}{257}-\frac{148}{771}i & -\frac{22}{257}-\frac{28}{771}i \\ 0 & 1 & 0 & 0 & \frac{185}{771}-\frac{124}{771}i & \frac{64}{257}+\frac{245}{771}i & -\frac{274}{771}-\frac{272}{771}i & -\frac{31}{771}-\frac{239}{771}i \\ 0 & 0 & 1 & 0 & \frac{100}{257}-\frac{83}{771}i & -\frac{29}{771}+\frac{50}{771}i & \frac{28}{771}-\frac{22}{257}i & \frac{172}{771}-\frac{25}{257}i \\ 0 & 0 & 0 & 1 & -\frac{200}{771}-\frac{116}{771}i & \frac{35}{257}-\frac{119}{771}i & \frac{67}{771}+\frac{44}{771}i & -\frac{29}{771}+\frac{50}{771}i \end{array}\right].$$

Con lo que

$$A^{-1}=\begin{bmatrix} \frac{112}{771}-\frac{88}{257}i & \frac{44}{771}-\frac{67}{771}i & \frac{104}{257}-\frac{148}{771}i & -\frac{22}{257}-\frac{28}{771}i \\ \frac{185}{771}-\frac{124}{771}i & \frac{64}{257}+\frac{245}{771}i & -\frac{274}{771}-\frac{272}{771}i & -\frac{31}{771}-\frac{239}{771}i \\ \frac{100}{257}-\frac{83}{771}i & -\frac{29}{771}+\frac{50}{771}i & \frac{28}{771}-\frac{22}{257}i & \frac{172}{771}-\frac{25}{257}i \\ -\frac{200}{771}-\frac{116}{771}i & \frac{35}{257}-\frac{119}{771}i & \frac{67}{771}+\frac{44}{771}i & -\frac{29}{771}+\frac{50}{771}i \end{bmatrix}.$$ ✔

Determinantes, adjunta y regla de Cramer

15 Para la matriz $A = \begin{bmatrix} -1 & 1 & 2 & 3 \\ 0 & 1 & -1 & 2 \\ 3 & 0 & -1 & 1 \\ 4 & -5 & 1 & 2 \end{bmatrix}$, hallar los cofactores *(a)* c_{11}, *(b)* c_{31} y *(c)* c_{41}.

Solución *(a)*

$$\begin{aligned} c_{11} &= (-1)^{1+1}\begin{vmatrix} 1 & -1 & 2 \\ 0 & -1 & 1 \\ -5 & 1 & 2 \end{vmatrix} \\ &= (1)(-1)^{1+1}M_{11} + (0)(-1)^{2+1}M_{21} + (-5)(-1)^{3+1}M_{31} \\ &= \begin{vmatrix} -1 & 1 \\ 1 & 2 \end{vmatrix} - 5\begin{vmatrix} -1 & 2 \\ -1 & 1 \end{vmatrix} \\ &= -3 - 5 = -8. \end{aligned}$$

(b)

$$\begin{aligned} c_{31} &= (-1)^{3+1}\begin{vmatrix} 1 & 2 & 3 \\ 1 & -1 & 2 \\ -5 & 1 & 2 \end{vmatrix} \\ &= (1)(-1)^{1+1}\begin{vmatrix} -1 & 2 \\ 1 & 2 \end{vmatrix} + (1)(-1)^{2+1}\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} \\ &\quad + (-5)(-1)^{3+1}\begin{vmatrix} 2 & 3 \\ -1 & 2 \end{vmatrix} = -40. \end{aligned}$$

(c)

$$\begin{aligned} c_{41} &= (-1)^{4+1}\begin{vmatrix} 1 & 2 & 3 \\ 1 & -1 & 2 \\ 0 & -1 & 1 \end{vmatrix} \\ &= -\left((-1)^{1+1}(1)\begin{vmatrix} -1 & 2 \\ -1 & 1 \end{vmatrix} + (-1)^{2+1}(1)\begin{vmatrix} 2 & 3 \\ -1 & 1 \end{vmatrix}\right) = 4. \end{aligned}$$ ✔

16 Utilizar los cofactores de la matriz A calculados en el ejercicio anterior para hallar el determinante de esa matriz.

Solución El determinante se puede calcular desarrollando por cofactores en la primera columna:

$$\begin{aligned} |A| &= a_{11}c_{11} + a_{21}c_{21} + a_{31}c_{31} + a_{41}c_{41} \\ &= (-1)(-8) + 0c_{21} + (3)(-40) + (4)(4) \\ &= -96. \end{aligned}$$ ✔

17 Calcular el determinante de la matriz A desarrollando por cofactores en una fila o columna que minimice los cálculos si

$$A = \begin{bmatrix} 2 & -1 & 3 & 0 & 2 \\ 0 & 2 & 2 & 1 & 1 \\ 1 & 0 & -1 & 0 & 1 \\ 3 & 0 & 2 & 0 & 1 \\ -0 & 4 & 1 & 2 & 0 \end{bmatrix}.$$

Solución Se desarrolla por cofactores por la columna o fila que tenga más ceros:

$$\begin{vmatrix} 2 & -1 & 3 & 0 & 2 \\ 0 & 2 & 2 & 1 & 1 \\ 1 & 0 & -1 & 0 & 1 \\ 3 & 0 & 2 & 0 & 1 \\ 0 & 4 & 1 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 2 & -1 & 3 & 2 \\ 1 & 0 & -1 & 1 \\ 3 & 0 & 2 & 1 \\ 0 & 4 & 1 & 0 \end{vmatrix} - 2 \begin{vmatrix} 2 & -1 & 3 & 2 \\ 0 & 2 & 2 & 1 \\ 1 & 0 & -1 & 1 \\ 3 & 0 & 2 & 1 \end{vmatrix}$$

$$= -(-1)\begin{vmatrix} 1 & -1 & 1 \\ 3 & 2 & 1 \\ 0 & 1 & 0 \end{vmatrix} + 4 \begin{vmatrix} 2 & 3 & 2 \\ 1 & -1 & 1 \\ 3 & 2 & 1 \end{vmatrix}$$

$$-2\left(-(-1)\begin{vmatrix} 0 & 2 & 1 \\ 1 & -1 & 1 \\ 3 & 2 & 1 \end{vmatrix} + (2)\begin{vmatrix} 2 & 3 & 2 \\ 1 & -1 & 1 \\ 3 & 2 & 1 \end{vmatrix} \right)$$

$$= -(1)\begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} + -2\left(-2\begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} + \begin{vmatrix} 1 & -1 \\ 3 & 2 \end{vmatrix} \right)$$

$$= -16. \checkmark$$

18 Calcular el determinante de la matriz A del ejercicio resuelto 17 utilizando las propiedades del inciso 5 del teorema 2.8 para "hacer ceros los elementos en la primera columna" hasta que se obtenga un determinante 3×3 y entonces calcularlo por cofactores o por la regla de Sarrus.

Solución

$$\begin{vmatrix} 2 & -1 & 3 & 0 & 2 \\ 0 & 2 & 2 & 1 & 1 \\ 1 & 0 & -1 & 0 & 1 \\ 3 & 0 & 2 & 0 & 1 \\ 0 & 4 & 1 & 2 & 0 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 2 & 2 & 1 & 1 \\ 2 & -1 & 3 & 0 & 2 \\ 3 & 0 & 2 & 0 & 1 \\ 0 & 4 & 1 & 2 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 2 & 2 & 1 & 1 \\ 0 & -1 & 5 & 0 & 0 \\ 3 & 0 & 2 & 0 & 1 \\ 0 & 4 & 1 & 2 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 2 & 2 & 1 & 1 \\ 0 & -1 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 & -2 \\ 0 & 4 & 1 & 2 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 2 & 2 & 1 & 1 \\ -1 & 5 & 0 & 0 \\ 0 & 5 & 0 & -2 \\ 4 & 1 & 2 & 0 \end{vmatrix}$$

$$= \begin{vmatrix} -1 & 5 & 0 & 0 \\ 2 & 2 & 1 & 1 \\ 0 & 5 & 0 & -2 \\ 4 & 1 & 2 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & -5 & 0 & 0 \\ 2 & 2 & 1 & 1 \\ 0 & 5 & 0 & -2 \\ 4 & 1 & 2 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & -5 & 0 & 0 \\ 0 & 12 & 1 & 1 \\ 0 & 5 & 0 & -2 \\ 4 & 1 & 2 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & -5 & 0 & 0 \\ 0 & 12 & 1 & 1 \\ 0 & 5 & 0 & -2 \\ 0 & 21 & 2 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 12 & 1 & 1 \\ 5 & 0 & -2 \\ 21 & 2 & 0 \end{vmatrix}$$

$$= -\left(-5\begin{vmatrix} 1 & 1 \\ 2 & 0 \end{vmatrix} + 2\begin{vmatrix} 12 & 1 \\ 21 & 2 \end{vmatrix}\right) = -16. \checkmark$$

19 Calcular el determinante de la matriz A del ejercicio 17 llevando esta matriz a forma triangular superior utilizando la propiedades 5 y 4 del teorema 2.8 (cfr. pág. 80).

Solución Por el ejercicio precedente se tiene

$$\begin{vmatrix} 2 & -1 & 3 & 0 & 2 \\ 0 & 2 & 2 & 1 & 1 \\ 1 & 0 & -1 & 0 & 1 \\ 3 & 0 & 2 & 0 & 1 \\ 0 & 4 & 1 & 2 & 0 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 2 & 2 & 1 & 1 \\ 0 & -1 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 & -2 \\ 0 & 4 & 1 & 2 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 \\ 0 & 0 & 5 & 0 & -2 \\ 0 & 4 & 1 & 2 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 12 & 1 & 1 \\ 0 & 0 & 5 & 0 & -2 \\ 0 & 0 & 21 & 2 & 0 \end{vmatrix}$$

$$= -12\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{12} & \frac{1}{12} \\ 0 & 0 & 5 & 0 & -2 \\ 0 & 0 & 21 & 2 & 0 \end{vmatrix}$$

$$= -12\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{12} & \frac{1}{12} \\ 0 & 0 & 0 & -\frac{5}{12} & -\frac{29}{12} \\ 0 & 0 & 0 & \frac{1}{4} & -\frac{21}{12} \end{vmatrix}$$

$$= (-12)\left(-\tfrac{5}{12}\right)\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{12} & \frac{1}{12} \\ 0 & 0 & 0 & 1 & \frac{29}{5} \\ 0 & 0 & 0 & \frac{1}{4} & -\frac{21}{12} \end{vmatrix}$$

$$= 5\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{12} & \frac{1}{12} \\ 0 & 0 & 0 & 1 & \frac{29}{5} \\ 0 & 0 & 0 & \frac{1}{4} & -\frac{21}{12} \end{vmatrix}$$

$$= 5\left(\tfrac{1}{4}\right)\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{12} & \frac{1}{12} \\ 0 & 0 & 0 & 1 & \frac{29}{5} \\ 0 & 0 & 0 & 1 & -7 \end{vmatrix}$$

$$= \tfrac{5}{4}\begin{vmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{12} & \frac{1}{12} \\ 0 & 0 & 0 & 1 & \frac{29}{5} \\ 0 & 0 & 0 & 0 & -\frac{64}{5} \end{vmatrix}$$

$$= \frac{5}{4}\left(-\frac{64}{5}\right) = -16. \checkmark$$

20 Calcular el determinante de la siguiente matriz llevando la misma a forma triangular superior utilizando las propiedades 5 y 4 del teorema 2.8 (cfr. pág. 80).

$$A = \begin{bmatrix} 1 & -2 & 3 & 4 & -5 \\ 2 & -3 & 2 & 1 & -4 \\ -3 & 4 & -6 & 1 & 3 \\ 5 & -8 & 12 & 15 & -20 \\ 7 & -18 & 16 & 12 & -30 \end{bmatrix}.$$

Solución

$$\begin{vmatrix} 1 & -2 & 3 & 4 & -5 \\ 2 & -3 & 2 & 1 & -4 \\ -3 & 4 & -6 & 1 & 3 \\ 5 & -8 & 12 & 15 & -20 \\ 7 & -18 & 16 & 12 & -30 \end{vmatrix} = \begin{vmatrix} 1 & -2 & 3 & 4 & -5 \\ 0 & 1 & -4 & -7 & 6 \\ 0 & -2 & 3 & 13 & -12 \\ 0 & 2 & -3 & -5 & 5 \\ 0 & -4 & -5 & -16 & 5 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & -2 & 3 & 4 & -5 \\ 0 & 1 & -4 & -7 & 6 \\ 0 & 0 & -5 & -1 & 0 \\ 0 & 0 & 5 & 9 & -7 \\ 0 & 0 & -21 & -44 & 29 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & -2 & 3 & 4 & -5 \\ 0 & 1 & -4 & -7 & 6 \\ 0 & 0 & -5 & -1 & 0 \\ 0 & 0 & 5 & 9 & -7 \\ 0 & 0 & -1 & -8 & 1 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & -2 & 3 & 4 & -5 \\ 0 & 1 & -4 & -7 & 6 \\ 0 & 0 & -1 & -8 & 1 \\ 0 & 0 & 5 & 9 & -7 \\ 0 & 0 & -5 & -1 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & -2 & 3 & 4 & -5 \\ 0 & 1 & -4 & -7 & 6 \\ 0 & 0 & -1 & -8 & 1 \\ 0 & 0 & 0 & -31 & -2 \\ 0 & 0 & 0 & 39 & -5 \end{vmatrix}$$

$$= \frac{1}{31}\begin{vmatrix} 1 & -2 & 3 & 4 & -5 \\ 0 & 1 & -4 & -7 & 6 \\ 0 & 0 & -1 & -8 & 1 \\ 0 & 0 & 0 & 1 & \frac{2}{31} \\ 0 & 0 & 0 & 39 & -5 \end{vmatrix}$$

$$= \frac{1}{31}\begin{vmatrix} 1 & -2 & 3 & 4 & -5 \\ 0 & 1 & -4 & -7 & 6 \\ 0 & 0 & -1 & -8 & 1 \\ 0 & 0 & 0 & 1 & \frac{2}{31} \\ 0 & 0 & 0 & 0 & -\frac{233}{31} \end{vmatrix} = 233. \checkmark$$

21 Calcular la matriz adjunta de la matriz

$$A = \begin{bmatrix} 1 & -i & 1+i \\ -i & 2i & 2 \\ -1 & 1 & 2+3i \end{bmatrix}.$$

Solución La matriz de signos correspondiente es $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$; así que

$$\operatorname{Cof}(A) = \begin{bmatrix} \begin{vmatrix} 2i & 2 \\ 1 & 2+3i \end{vmatrix} & -\begin{vmatrix} -i & 2 \\ -1 & 2+3i \end{vmatrix} & \begin{vmatrix} -i & 2i \\ -1 & 1 \end{vmatrix} \\ -\begin{vmatrix} -i & 1+i \\ 1 & 2+3i \end{vmatrix} & \begin{vmatrix} 1 & 1+i \\ -1 & 2+3i \end{vmatrix} & -\begin{vmatrix} 1 & -i \\ -1 & 1 \end{vmatrix} \\ \begin{vmatrix} -i & 1+i \\ 2i & 2 \end{vmatrix} & -\begin{vmatrix} 1 & 1+i \\ -i & 2 \end{vmatrix} & \begin{vmatrix} 1 & -i \\ -i & 2i \end{vmatrix} \end{bmatrix} =$$

$$= \begin{bmatrix} -8+4i & -5+2i & i \\ -2+3i & 3+4i & -1+i \\ 2-4i & -1-i & 1+2i \end{bmatrix};$$

por tanto,

$$\operatorname{Adj}(A)=\begin{bmatrix}-8+4i & -2+3i & 2-4i\\ -5+2i & 3+4i & -1-i\\ i & -1+i & 1+2i\end{bmatrix}. \checkmark$$

22 Calcular la inversa de la matriz A del ejercicio anterior por medio de la adjunta.

Solución Al desarrollar por cofactores en la tercera fila de A se tiene

$$\begin{aligned}\det(A) &= (-1)(2-4i)+(1)(-1-i)+(2+3i)(1+2i)\\ &= -7+10i.\end{aligned}$$

Entonces,

$$\begin{aligned}A^{-1} &= \frac{1}{\det(A)}\operatorname{Adj}(A)\\ &= \frac{1}{-7+10i}\begin{bmatrix}-8+4i & -2+3i & 2-4i\\ -5+2i & 3+4i & -1-i\\ i & -1+i & 1+2i\end{bmatrix}\\ &= \begin{bmatrix}\frac{96}{149}+\frac{52}{149}i & \frac{44}{149}-\frac{1}{149}i & -\frac{54}{149}+\frac{8}{149}i\\ \frac{55}{149}+\frac{36}{149}i & \frac{19}{149}-\frac{58}{149}i & -\frac{3}{149}+\frac{17}{149}i\\ \frac{10}{149}-\frac{7}{149}i & \frac{17}{149}+\frac{3}{149}i & \frac{13}{149}-\frac{24}{149}i\end{bmatrix}. \checkmark\end{aligned}$$

23 Calcular el determinante de la matriz A utilizando las propiedades 5 y 4 del teorema 2.8 (cfr. pág. 80) y desarrollar por cofactores o regla de Sarrus si

$$A=\begin{bmatrix}1 & -1 & 2 & i & 1-i\\ -i & 2+i & 3 & -1 & 2\\ -3 & 1 & -1 & 2 & i\\ -1+i & -1 & 1 & 2 & -3i\\ 2 & 2i & -i & 2+3i & i\end{bmatrix}.$$

Solución

$$\begin{aligned}\begin{vmatrix}1 & -1 & 2 & i & 1-i\\ -i & 2+i & 3 & -1 & 2\\ -3 & 1 & -1 & 2 & i\\ -1+i & -1 & 1 & 2 & -3i\\ 2 & 2i & -i & 2+3i & i\end{vmatrix} &= \begin{vmatrix}1 & -1 & 2 & i & 1-i\\ 0 & 2 & 3+2i & -2 & 3+i\\ 0 & -2 & 5 & 2+3i & 3-2i\\ 0 & -2+i & 3-2i & 3+i & -5i\\ 0 & 2+2i & -4-i & 2+i & -2+3i\end{vmatrix}\\ &= \begin{vmatrix}1 & -1 & 2 & i & 1-i\\ 0 & 2 & 3+2i & -2 & 3+i\\ 0 & 0 & 8+2i & 3i & 6-i\\ 0 & i & 6 & 1+i & 3-4i\\ 0 & 3i & -1-3i & 5+2i & -2-2i\end{vmatrix}\\ &= i\begin{vmatrix}1 & -1 & 2 & i & 1-i\\ 0 & 2 & 3+2i & -2 & 3+i\\ 0 & 0 & 8+2i & 3i & 6-i\\ 0 & 1 & -6i & 1-i & -4-3i\\ 0 & 3i & -1-3i & 5+2i & -2-2i\end{vmatrix}\end{aligned}$$

$$= -i\begin{vmatrix} 1 & -1 & 2 & i & 1-i \\ 0 & 1 & -6i & 1-i & -4-3i \\ 0 & 0 & 8+2i & 3i & 6-i \\ 0 & 2 & 3+2i & -2 & 3+i \\ 0 & 3i & -1-3i & 5+2i & -2-2i \end{vmatrix}$$

$$= -i\begin{vmatrix} 1 & -1 & 2 & i & 1-i \\ 0 & 1 & -6i & 1-i & -4-3i \\ 0 & 0 & 8+2i & 3i & 6-i \\ 0 & 0 & 3+14i & -4+2i & 11+7i \\ 0 & 0 & -19-3i & 2-i & -11+10i \end{vmatrix}$$

$$= -i\begin{vmatrix} 8+2i & 3i & 6-i \\ 3+14i & -4+2i & 11+7i \\ -19-3i & 2-i & -11+10i \end{vmatrix}$$

$$= -i\left\{\begin{array}{c} \left(\begin{array}{c} (8+2i)(-4+2i)(-11+10i) \\ +(3+14i)(2-i)(6-i) \\ +(-19-3i)(11+7i)(3i) \end{array}\right) \\ -\left(\begin{array}{c} (6-i)(-4+2i)(-19-3i) \\ +(11+7i)(2-i)(8+2i) \\ +(-11+10i)(3+14i)(3i) \end{array}\right) \end{array}\right\}$$

$$= -i((959-882i)-(1064-675i))$$
$$= -207+105i. \checkmark$$

24 Calcular la solución del sistema

$$\begin{aligned} (1+2i)x-iy+z &= 2+i \\ 2x+(3-i)y+iz &= 5 \\ (-2-i)x+(2+i)y-2z &= -2 \end{aligned}$$

mediante la regla de Cramer.

Solución

$$\Delta = \begin{vmatrix} 1+2i & -i & 1 \\ 2 & 3-i & i \\ -2-i & 2+i & -2 \end{vmatrix}$$

$$= \left\{\left(\begin{array}{c} -2(1+2i)(3-i)+2(2+i)+(-2-i)(i)(-i) \\ -((3-i)(-2-i)+i(2+i)(1+2i)+4i) \end{array}\right)\right\}$$

$$= -8-9i-(-12+3i)$$
$$= 4-12i;$$

$$\Delta_1 = \begin{vmatrix} 2+i & -i & 1 \\ 5 & 3-i & i \\ -2 & 2+i & -2 \end{vmatrix}$$

$$= (-2(2+i)(3-i)+10+5i-2)-\big((-6+2i)+(2+i)^2 i+10i\big)$$
$$= -6+3i-(-10+15i)$$
$$= 4-12i;$$

$$\begin{aligned}
\Delta_2 &= \begin{vmatrix} 1+2i & 2+i & 1 \\ 2 & 5 & i \\ -2-i & -2 & -2 \end{vmatrix} \\
&= \left(-10-20i-4-i(2+i)^2\right) \\
&\quad -(-10-5i+4-2i-8-4i) \\
&= -10-23i-(-14-11i) \\
&= 4-12i;
\end{aligned}$$

$$\begin{aligned}
\Delta_3 &= \begin{vmatrix} 1+2i & -i & 2+i \\ 2 & 3-i & 5 \\ -2-i & 2+i & -2 \end{vmatrix} \\
&= \left((-2-4i)(3-i)+2(2+i)^2-5+10i\right)-\left(-(2+i)^2(3-i)+5(1+2i)(2+i)+4i\right) \\
&= -9+8i-(-13+20i) \\
&= 4-12i.
\end{aligned}$$

Por tanto,
$$x = \frac{\Delta_1}{\Delta} = 1,$$
$$y = \frac{\Delta_1}{\Delta} = 1,$$
$$z = \frac{\Delta_1}{\Delta} = 1. \quad ✔$$

2.3.2 Ejercicios propuestos

El lector encontrará la respuesta a los ejercicios en cursiva en el apéndice E al final del libro.

Matrices invertibles y sus inversas (respuestas en páginas 1075-1076)

En los ejercicios 1 a 7 encontrar, si existe, la inversa de la matriz dada mediante el método de Gauss-Jordan.

1 $\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$.

2 $\begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}$.

3 $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$.

4 $\begin{bmatrix} 2 & 3 & 0 \\ 1 & 2 & -1 \\ 4 & 5 & 1 \end{bmatrix}$.

5 $\begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 3 \\ -3 & 1 & -7 \end{bmatrix}$.

6 $\begin{bmatrix} -1 & 2 & 3 \\ 2 & -3 & -1 \\ 1 & -1 & 1 \end{bmatrix}$.

7 $\begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$.

8 Escriba las matrices de los ejercicios 1 a 7, donde sea posible, como producto de matrices elementales.

En los ejercicios 9 a 13 utilizar las matrices inversas encontradas en los problemas 1, 2, 3, 4 y 6 para resolver los sistemas lineales.

9 $\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$.

10 $\begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ -4 \end{bmatrix}$.

11 $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \\ -4 \end{bmatrix}$.

12 $\begin{bmatrix} 2 & 3 & 0 \\ 1 & 2 & -1 \\ 4 & 5 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -3 \\ -4 \\ -2 \end{bmatrix}$.

13 $\begin{bmatrix} -1 & 2 & 3 \\ 2 & -3 & -1 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$.

14 Demostrar que una matriz $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ es invertible si y sólo si $ad - bc \neq 0$ y que en ese caso

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

(Sugerencia: Utilizar el método de Gauss-Jordan.)

Utilizar el ejercicio 14 para determinar si las matrices de los ejercicios 15 a 18 son invertibles y calcular en tal caso las matrices inversas.

15 $\begin{bmatrix} -1 & 2 \\ 3 & -8 \end{bmatrix}$.

16 $\begin{bmatrix} 2 & -3 \\ 5 & 7 \end{bmatrix}$.

17 $\begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix}$.

18 $\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$, $\alpha, \beta \in \mathbb{R} - \{0\}$.

En los ejercicios 19 a 37, encontrar la inversa de la matriz dada (si existe) por el método de Gauss-Jordan.

19 $\begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -2 \\ 3 & -3 & 5 \end{bmatrix}$.

20 $\begin{bmatrix} 1 & -2 & -1 \\ -2 & 3 & 2 \\ -2 & 1 & 1 \end{bmatrix}$.

21 $\begin{bmatrix} -2 & 5 & 3 \\ 3 & -1 & -2 \\ -1 & 1 & 1 \end{bmatrix}$.

22 $\begin{bmatrix} 7 & 40 & 19 \\ -4 & -26 & -13 \\ -3 & -17 & -8 \end{bmatrix}$.

23 $\begin{bmatrix} 2 & -1 & 3 \\ 3 & -1 & 4 \\ 2 & -1 & 1 \end{bmatrix}$.

24 $\begin{bmatrix} 5 & -3 & 2 \\ -3 & 6 & 3 \\ 4 & -1 & 3 \end{bmatrix}$.

25 $\begin{bmatrix} 1 & -2 & 4 \\ 5 & -10 & 17 \\ -3 & 5 & -12 \end{bmatrix}$.

26 $\begin{bmatrix} 3 & 1 & 0 \\ 1 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$.

27 $\begin{bmatrix} 1 & -2 & 1 & 3 \\ -2 & 5 & -2 & -5 \\ 3 & -6 & 4 & 9 \\ 4 & -8 & 4 & 11 \end{bmatrix}$.

28 $\begin{bmatrix} 2 & -3 & 1 & 2 \\ 1 & 2 & 3 & 3 \\ 3 & -1 & 4 & 5 \\ -1 & 2 & -1 & 4 \end{bmatrix}$.

29 $\begin{bmatrix} -1 & 2 & -1 & 1 \\ 3 & -5 & 3 & -3 \\ -2 & 4 & -3 & 2 \\ 2 & -4 & -2 & -1 \end{bmatrix}$.

30 $\begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 6 & 1 & 6 \\ 4 & 2 & 6 & 12 \\ 7 & 14 & 7 & 20 \end{bmatrix}$.

31 $\begin{bmatrix} 1 & -1 & 2 & -3 \\ 2 & -2 & 1 & -1 \\ -1 & 9 & 1 & 1 \\ 0 & -1 & 1 & -2 \end{bmatrix}$.

32 $\begin{bmatrix} 1 & -1 & 2 & -3 \\ 2 & -2 & 1 & -1 \\ -1 & -25 & 1 & 1 \\ -2 & -1 & 1 & -2 \end{bmatrix}$.

33 $\begin{bmatrix} 1 & 1 & 1 & 5 \\ 1 & 3 & 2 & 9 \\ 1 & 1 & -1 & 1 \\ 1 & 2 & 1 & 6 \end{bmatrix}$.

34 $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 3 \\ 1 & -1 & 2 & 1 \\ 1 & 2 & 1 & 2 \end{bmatrix}$.

35 $\begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 2 & -1 & 4 & 6 & 8 \\ -2 & 2 & -3 & -6 & -8 \\ 3 & -3 & 6 & 8 & 12 \\ 5 & -5 & 10 & 15 & 19 \end{bmatrix}$.

36 $\begin{bmatrix} 1 & -2 & 1 & 1 & -1 \\ 3 & -5 & 3 & 4 & -2 \\ -2 & -4 & -1 & -3 & 3 \\ -1 & 2 & -1 & -2 & 2 \\ 2 & -4 & 2 & 2 & -3 \end{bmatrix}$.

37 $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & -2 & 0 & 0 & 0 \\ 0 & 3 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & -3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{bmatrix}$.

38 Sea

$$A^{-1} = \begin{bmatrix} 1 & -2 & 1 \\ 3 & -2 & 0 \\ 0 & 2 & 1 \end{bmatrix}.$$

Hallar, si es posible, una matriz C tal que

$$AC = \begin{bmatrix} 1 & 1 \\ -1 & 2 \\ 2 & -1 \end{bmatrix}.$$

39 Sea $A^{-1} = \begin{bmatrix} 2 & -1 & 1 \\ -2 & 1 & 2 \\ 1 & 1 & -2 \end{bmatrix}$. Encontrar, si es posible, una matriz B tal que

$$ABA = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & 3 \\ 0 & 1 & -1 \end{bmatrix}.$$

40 Sea $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -1 & 0 \\ 2 & 1 & 0 \end{bmatrix}$. Hallar, si es que existe, una matriz B tal que $A^2 - A = AB$.

41 Sea $A = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}$. Resolver la ecuación matricial

$$A^2X - A = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}.$$

42 Encontrar los valores de α tales que la matriz

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & \alpha & -1 \\ 3 & -2 & 2 \end{bmatrix}$$

sea invertible.

43 Hallar los valores de α tales que la matriz

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & \alpha \\ -1 & -1 & 1 \end{bmatrix}$$

sea invertible.

44 Mostrar que dos matrices $A, B \in \mathfrak{M}_{m\times n}$ son equivalentes si y sólo si existen matrices elementales $E_1, E_2, \ldots, E_k$ de orden m tales que

$$B = E_k \cdots E_2 E_1 A.$$

45 Demostrar que dos matrices A, B de tamaño $m \times n$ son equivalentes si y sólo si existe una matriz cuadrada, C, de orden m e invertible tal que $B = CA$.

46 Si una matriz cuadrada satisface $A^2 = \mathscr{O}$, demostrar que entonces $A - I$ es una matriz invertible. ¿Cuál es su inversa? (Donde I es la matriz identidad).

Los ejercicios 47 a 58 contienen afirmaciones que son falsas o verdaderas. Si la afirmación es verdadera se debe demostrar con rigor su validez y si es falsa se tiene que exhibir un contraejemplo para mostrar que no es cierta. En cada caso se supone que las matrices involucradas tienen los tamaños adecuados para efectuar las correspondientes operaciones.

47 C invertible, $AC = BC \Rightarrow A = B$.

48 C invertible, $AC = CB \Rightarrow A = B$.

49 C invertible, $AC = \mathscr{O} \Rightarrow A = \mathscr{O}$.

50 Si $AC = B$ y dos de estas matrices son invertibles, entonces la tercera también es invertible.

51 Si $AC = B$ y dos de estas matrices son singulares (no invertibles), entonces la tercera también es singular.

52 A invertible $\Rightarrow A^2$ invertible.

53 A^2 invertible $\Rightarrow A$ invertible.

54 A^3 invertible $\Rightarrow A$ invertible.

55 A, B invertibles $\Rightarrow A + B$ invertible.

56 A invertible $\Rightarrow A + A$ invertible.

57 A invertible $\Rightarrow A + A^t$ invertible.

58 A, B no singulares (invertibles) $\Rightarrow A+B$ no singular.

59 Una matriz cuadrada A tiene potencia nula si $A^k = \mathscr{O}$ para algún entero positivo k.

(a) Dar un ejemplo de una matriz no cero de orden 2 que tenga potencia nula.

(b) Probar que toda matriz invertible no puede tener potencia nula.

60 Si A es una matriz cuadrada y $A^2 = A$, probar que si A es invertible entonces $A = I$.

En los ejercicios 61 a 62 determinar si la matriz es invertible y en tal caso calcular la inversa mediante el método de Gauss-Jordan.

61 $\begin{bmatrix} 1 & 1+i & -1 & 2 \\ 2 & 2+i & -2 & 4 \\ -1+i & -2 & 1 & -2+2i \\ i & -1+i & -i & i \end{bmatrix}$.

62 $\begin{bmatrix} 6+2i & 13+3i & -8+3i & -2+4i \\ 13+3i & 27+4i & -16+8i & -3+9i \\ -8+3i & -16+8i & 5-10i & -1-6i \\ -2+4i & -3+9i & -1-6i & -3-2i \end{bmatrix}$.

Determinantes, adjunta y regla de Cramer (respuestas en páginas 1076-1077)

En los ejercicios 63 a 66, calcular los determinantes por medio de la regla de Sarrus (cfr. (2.7), pág. 78).

63 $\begin{vmatrix} 2 & -1 & 3 \\ 4 & -3 & 2 \\ 1 & 4 & 6 \end{vmatrix}$.

64 $\begin{vmatrix} 3 & 1 & 1 \\ -2 & 1 & 2 \\ 2 & -1 & 3 \end{vmatrix}$.

65 $\begin{vmatrix} 2 & 1 & 1 \\ 3 & -2 & 4 \\ 5 & -7 & 2 \end{vmatrix}$.

66 $\begin{vmatrix} -5 & -5 & -3 \\ -3 & 7 & 5 \\ 7 & 5 & 4 \end{vmatrix}$.

En los ejercicios 67 y 68, calcular los cofactores indicados para la matriz A.

67 $A = \begin{bmatrix} 3 & 7 & -5 & 4 \\ -8 & -3 & 2 & 3 \\ -2 & 7 & 6 & 4 \\ -5 & 9 & -1 & -0 \end{bmatrix}$. *(a)* c_{13}, *(b)* c_{31}, *(c)* c_{11}.

68 $A = \begin{bmatrix} -1 & 2 & 3 & 1 \\ 2 & -1 & 1 & 1 \\ 3 & -2 & 1 & -3 \\ 2 & 1 & 1 & 4 \end{bmatrix}$. *(a)* c_{43}, *(b)* c_{22}, *(c)* c_{14}.

69 Escribir la generalización natural de la regla de Sarrus para una matriz 4×4; calcular con esta regla el determinante de la matriz

$$A = \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & 2 & 3 \\ -1 & 2 & -1 & 1 \\ 0 & 1 & -1 & 3 \end{bmatrix}$$

y encontrar el determinante por medio de cofactores. ¿Se obtiene el mismo resultado?

En los ejercicios 70 a 77, calcular el determinante por medio de cofactores.

70 $\begin{vmatrix} 2 & 0 & -1 & 3 \\ -1 & 3 & 1 & 4 \\ 1 & 0 & -1 & 1 \\ 2 & 2 & 1 & -1 \end{vmatrix}.$

71 $\begin{vmatrix} 3 & 0 & 1 & 0 \\ -1 & 1 & 3 & 5 \\ 2 & 1 & 0 & 4 \\ 1 & 1 & 2 & 1 \end{vmatrix}.$

72 $\begin{vmatrix} -4 & 2 & 1 & 3 \\ 0 & -1 & 1 & 1 \\ 1 & 2 & 1 & -1 \\ 0 & 1 & 2 & 3 \end{vmatrix}.$

73 $\begin{vmatrix} 1 & 2 & -1 & 1 \\ 2 & -3 & 2 & -2 \\ 3 & -1 & 1 & -1 \\ 4 & -1 & 1 & 2 \end{vmatrix}.$

74 $\begin{vmatrix} 2 & -1 & 3 & -1 & 2 \\ 0 & 2 & 4 & -2 & 3 \\ 0 & 1 & 1 & 0 & 5 \\ 0 & -1 & 3 & 0 & 2 \\ 0 & 1 & -1 & 1 & 2 \end{vmatrix}.$

75 $\begin{vmatrix} -3 & 2 & -1 & -1 & -1 & 2 \\ 0 & 4 & 2 & 2 & 2 & 3 \\ 0 & 0 & 3 & 3 & 5 & 6 \\ 0 & 0 & 0 & 2 & -1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 5 & 3 & 2 \end{vmatrix}.$

76 $\begin{vmatrix} 1 & 0 & 3 & 2 & 0 \\ -2 & 1 & 0 & -1 & -2 \\ 3 & -1 & 6 & 1 & 0 \\ 4 & 2 & 0 & 3 & 2 \\ -7 & 3 & 0 & 2 & 1 \end{vmatrix}.$

77 $\begin{vmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 0 & 2 & 1 & 1 \end{vmatrix}.$

78 Las matrices de los ejercicios 75 y 77 tienen componentes nulas excepto en algunas de las entradas de una submatriz R de tamaño $r \times r$ y en las entradas de otra submatriz S de tamaño $s \times s$, cuyas diagonales están en la diagonal de la matriz original de tamaño $n \times n$ y donde, además, $r + s = n$. Demostrar que si este es el caso, entonces

$$\det(A) = \det(R)\det(S).$$

79 Establecer y demostrar un resultado análogo al del ejercicio anterior, pero suponiendo que las submatrices tienen sus contradiagonales en la contradiagonal de la matriz original.

80 Establecer y demostrar generalizaciones de los dos ejercicios precedentes cuando el número de submatrices es mayor a dos.

Utilizar los ejercicios 78, 79 y 80 para calcular los determinantes de las matrices contenidas en los ejercicios 81 y 82.

81 $\begin{bmatrix} 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & -3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \end{bmatrix}.$

82 $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 2 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

En los ejercicios 83 a 90, calcular el determinante utilizando la propiedad 5 del teorema 2.8; proceder como en los ejercicios resueltos 18 y 19 (cfr. pág. 96).

83 $\begin{vmatrix} 2 & 4 & 6 & -4 \\ 2 & -1 & 2 & 3 \\ -3 & 1 & 2 & 1 \\ -2 & 1 & 1 & 1 \end{vmatrix}$.

84 $\begin{vmatrix} 1 & -3 & 2 & 1 \\ -1 & 2 & 1 & 1 \\ 3 & -2 & 5 & 2 \\ 4 & -1 & 1 & 2 \end{vmatrix}$.

85 $\begin{vmatrix} 3 & 2 & 1 & 2 \\ 2 & -1 & 1 & 2 \\ 5 & -2 & 1 & 4 \\ -1 & 1 & 2 & 1 \end{vmatrix}$.

86 $\begin{vmatrix} 1 & -1 & 2 & 3 & 2 \\ -1 & 2 & 4 & -1 & 1 \\ 3 & -2 & 1 & 1 & 2 \\ 2 & -1 & 1 & 3 & 2 \\ 3 & 3 & -2 & 1 & 3 \end{vmatrix}$.

87 $\begin{vmatrix} 5 & 1 & -2 & -1 \\ 11 & 2 & 3 & 2 \\ 3 & 2 & 3 & 3 \\ 4 & 2 & 3 & -2 \end{vmatrix}$.

88 $\begin{vmatrix} 2 & -1 & 2 & -1 & 2 \\ 3 & 2 & 3 & 2 & 1 \\ -1 & 4 & 3 & 1 & 1 \\ 1 & -2 & 1 & 3 & 2 \\ 2 & 1 & 1 & -1 & 1 \end{vmatrix}$.

89 $\begin{vmatrix} 2 & 1 & -1 & 5 & 7 \\ 3 & 1 & 2 & -2 & -2 \\ 5 & 2 & 1 & 5 & 3 \\ 3 & -1 & 3 & 3 & -1 \\ 2 & 4 & 3 & 2 & 2 \end{vmatrix}$.

90 $\begin{vmatrix} 1 & -1 & 2 & 1 & 3 & 7 \\ -1 & 4 & -1 & -2 & -2 & -6 \\ 2 & -2 & 3 & 4 & 2 & 13 \\ -3 & 3 & -5 & 2 & 4 & 18 \\ 1 & -1 & 2 & 1 & 5 & 3 \\ -2 & 2 & -4 & -2 & -4 & 3 \end{vmatrix}$.

En los ejercicios 91 a 96, A es una matriz de tamaño 3×3 con $\det(A) = -4$.

91 Calcular $\det(A^2)$.

92 Encontrar $\det(4A)$.

93 Hallar $\det(A^{-1})$.

94 Calcular $\det(A^m)$, donde m es cualquier entero positivo.

95 Hallar $\det(A+A)$.

96 Calcular $\det(A^t)$.

97 Sean α_1, α_2 y α_3 números reales. Mostrar que

$$\begin{vmatrix} 1 & \alpha_1 & \alpha_1^2 \\ 1 & \alpha_2 & \alpha_2^2 \\ 1 & \alpha_3 & \alpha_3^2 \end{vmatrix} = (\alpha_2 - \alpha_1)(\alpha_3 - \alpha_1)(\alpha_3 - \alpha_2).$$

98 Sean α_i, $i = 1, \ldots, n$, números reales. Demostrar que

$$\begin{vmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{vmatrix} = \prod_{i<j}(\alpha_j - \alpha_i).$$

El símbolo $\prod_{i<j}(\alpha_j - \alpha_i)$ significa el producto de todos los posibles factores $(\alpha_j - \alpha_i)$ con $j > i$, cuando i, j varían entre 1 y n. A este valor se le conoce como el determinante de Vandermonde.

99 Sean A una matriz 3×3 y $\alpha \in \mathbb{R}$, probar que

$$\det(\alpha A) = \alpha^3 \det(A).$$

100 Sean A una matriz $n \times n$ y $\alpha \in \mathbb{R}$, probar que

$$\det(\alpha A) = \alpha^n \det(A).$$

101 Sea

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

mostrar utilizando el teorema 2.9, que A es invertible si y sólo si $ad - bc \neq 0$ y que en tal caso

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

102 Utilizar el ejercicio precedente para calcular (si existe):

(a) $\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}^{-1}$, *(b)* $\begin{bmatrix} 2 & 4 \\ -3 & 2 \end{bmatrix}^{-1}$ *(c)* $\begin{bmatrix} 3 & 6 \\ 1 & 2 \end{bmatrix}^{-1}$.

En los ejercicios 103 a 106, calcular la inversa, si existe, de la matriz dada utilizando el método de la adjunta (teorema 2.9, pág. 83).

103 $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & -2 & 2 \\ 2 & 1 & 1 \end{bmatrix}$.

104 $A = \begin{bmatrix} 3 & -1 & 2 \\ -2 & 1 & 2 \\ 3 & -1 & 4 \end{bmatrix}$.

105 $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 3 \\ 3 & 4 & 2 \end{bmatrix}$.

106 $A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 0 & 2 \\ 1 & -1 & 1 \end{bmatrix}$.

107 Si A es una matriz cuadrada singular (no invertible) y B es una matriz cuadrada del mismo orden, demostrar que AB es también una matriz singular.

108 Demostrar que una matriz cuadrada es invertible si y sólo si su adjunta también es una matriz invertible y que además se tiene, entonces, $(\text{Adj}(A))^{-1} = (1/|A|)A$.

109 Probar que si $A \in \mathfrak{M}_n$, entonces $|\text{Adj}(A)| = |A|^{n-1}$.

110 Mostrar que si A es una matriz cuadrada de orden $n > 1$, entonces $\mathrm{Adj}(\mathrm{Adj}(A)) = |A|^{n-2}A$.

En los ejercicios 111 a 114, determinar los valores de $\lambda \in \mathbb{R}$ que hacen que la matriz sea singular (no invertible).

111 $\begin{bmatrix} \lambda+3 & 1 \\ 2 & \lambda+2 \end{bmatrix}$.

112 $\begin{bmatrix} 1-\lambda & 3 & 2 \\ 0 & 2+\lambda & 1 \\ 0 & -1 & 2-\lambda \end{bmatrix}$.

113 $\begin{bmatrix} 2 & 1 & -1 \\ \lambda+2 & \lambda & 0 \\ \lambda-3 & 2 & 0 \end{bmatrix}$.

114 $\begin{bmatrix} 0 & 0 & 2-\lambda \\ 4-\lambda & 3 & -3 \\ 4 & -\lambda & 2 \end{bmatrix}$.

Resolver los sistemas de los ejercicios 115 a 117, mediante la regla de Cramer (teorema 2.11, pág. 84).

115 $\begin{aligned} x_1 - 3x_2 &= -7 \\ 2x_1 - x_2 &= 4 \end{aligned}$

116 $\begin{aligned} 2x_1 - 3x_2 + x_3 &= 4 \\ 4x_1 - 5x_2 + 3x_3 &= 4 \\ -3x_1 + 2x_2 - 3x_3 &= 1 \end{aligned}$

117 $\begin{aligned} x_1 - x_2 + x_3 &= -2 \\ 2x_1 - x_2 + 3x_3 &= -9 \\ 5x_1 - 3x_2 + 4x_3 &= -11 \end{aligned}$

118 Hallar la componente x_2 de la solución del sistema lineal.

$$\begin{aligned} x_1 + x_2 - 3x_3 + x_4 &= -2 \\ 2x_1 + x_2 + 2x_4 &= 3 \\ x_2 - 6x_3 - x_4 &= -8 \\ 3x_1 + x_2 + x_4 &= 3 \end{aligned}$$

119 Hallar la componente x_4 de la solución del sistema del lineal.

$$\begin{aligned} x_1 - x_2 + x_3 - x_4 &= 2 \\ x_1 + 2x_2 - x_4 &= -2 \\ x_1 - 3x_2 + 2x_4 &= 9 \\ -x_1 + 2x_2 - 3x_3 + x_4 &= -5 \end{aligned}$$

En los ejercicios 120 a 124, calcular el determinante de la matriz dada.

120 $\begin{bmatrix} i & 2+i & -i \\ 2 & 3+2i & 1+i \\ 2-i & 4+1 & 1 \end{bmatrix}$.

121 $\begin{bmatrix} -4i & 2 & 1 \\ 1+i & -2 & 1 \\ 3+2i & -2+i & 2 \end{bmatrix}$.

122 $\begin{bmatrix} 1 & 1+i & 2 & 1 \\ -1 & 1-i & 2 & 2i \\ 4i & -2i & 5 & 2i \\ -3+i & 2i & -4i & 1+6i \end{bmatrix}$.

123 $\begin{bmatrix} i & 0 & 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & i & 1+2i \\ 0 & 0 & 0 & 2 & -i & 4 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$.

124 $\begin{bmatrix} 1 & i & 2i & -3i & 4 \\ i & 3i & -2 & 3 & 4i \\ -2 & 2i & 1 & -2 & 1 \\ 1+i & 1-i & -1 & i & 3 \\ -2 & -2i & 1 & 2 & 1 \end{bmatrix}$.

Espacios vectoriales, producto interior, normas, valores y vectores propios

3 Espacios vectoriales

En este capítulo estudiaremos el siguiente concepto clave en álgebra lineal. Pero antes, motivaremos las ideas principales por medio de un repaso de conceptos elementales del plano cartesiano. Después extenderemos estos conceptos de manera natural a espacios de vectores más generales para, posteriormente, continuar con un estudio abstracto y completo de estos entes que llamaremos espacios vectoriales.

3.1 Geometría de los espacios $\mathbb{R}^n$

En esta sección el objetivo fundamental es generalizar las características geométricas esenciales que poseen los vectores en el plano cartesiano y en el espacio de tres dimensiones a espacios cuyos vectores tienen mayor número de coordenadas, los llamados espacios $\mathbb{R}^n$. Para ello comenzamos, en la primera subsección, con un repaso de estas características en el plano cartesiano. Bien pudimos emplear como modelo para este propósito el espacio tridimensional pero, por razones de sencillez en cuanto a los bosquejos geométricos, hemos preferido utilizar el plano cartesiano; sin embargo, como el lector podrá constatar fácilmente por sí mismo, todo lo que hagamos en el siguiente apartado para el plano cartesiano es completamente válido cuando se traslada al espacio de tres dimensiones.

3.1.1 *El plano cartesiano* $\mathbb{R}^2$

Definición 3.1 *Definimos*

$$\mathbb{R}^2 = \{(x,y) \,|\, x,y \in \mathbb{R}\}.$$

Geométricamente, $\mathbb{R}^2$ es el plano cartesiano con el que el lector está familiarizado de sus cursos elementales y que ilustramos en la figura 3-1.

Las características esenciales, algebraicas y geométricas, de $\mathbb{R}^2$ son:

1. **Igualdad**: Si $\vec{u} = (x_1, y_1), \vec{v} = (x_2, y_2) \in \mathbb{R}^2$, $\vec{u} = \vec{v} \Leftrightarrow x_1 = x_2$ y $y_1 = y_2$.
2. **Suma**: Si $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$,

$$\vec{u} + \vec{v} = (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2).$$

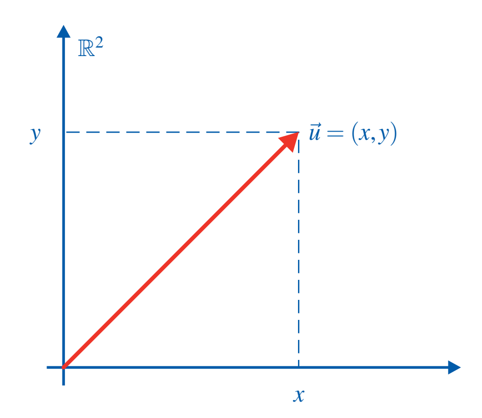


Figura 3-1 • $\mathbb{R}^2$, plano donde cada punto (vector) $\vec{u}$ se localiza mediante un par ordenado (x,y).

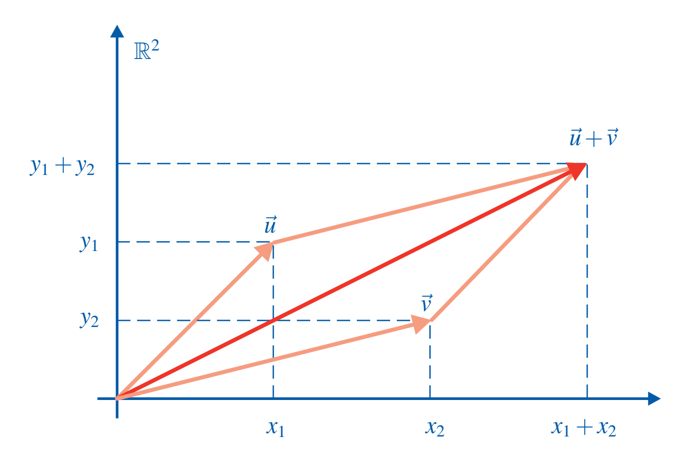


Figura 3-2 • La suma de dos vectores en el plano es la diagonal del paralelogramo que se genera a partir de esos vectores.

La suma de dos vectores de $\mathbb{R}^2$ se obtiene, geométricamente, por la diagonal del paralelogramo como se indica en la figura 3-2.

3. **Producto de un escalar por un vector**. Si $\lambda \in \mathbb{R}$ y $\vec{u}=(x,y)\in\mathbb{R}^2$, $\lambda\vec{u}=(\lambda x,\lambda y)$.

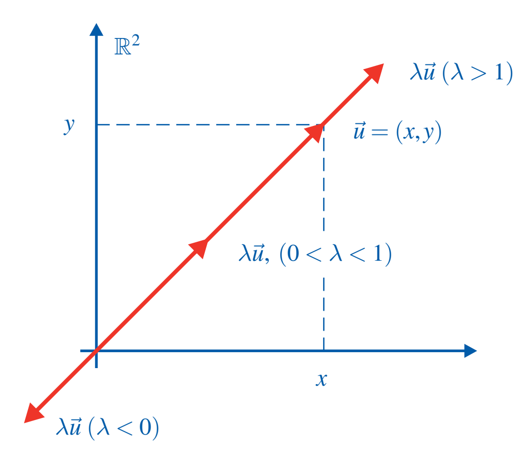


Figura 3-3 • Producto de un escalar por un vector.

Así, el vector $\lambda\vec{u}$ es un vector paralelo a $\vec{u}$ con un cambio de escala y/o de sentido, tal como queda ilustrado en la figura 3-3.

4. **Norma o magnitud de un vector**. Si $\vec{u} = (x, y)$, se define y denota la **norma** de $\vec{u}$ como

$$\|\vec{u}\| = \sqrt{x^2 + y^2}.$$

La norma representa la magnitud o la longitud del vector $\vec{u}$ (véase la figura 3-4).

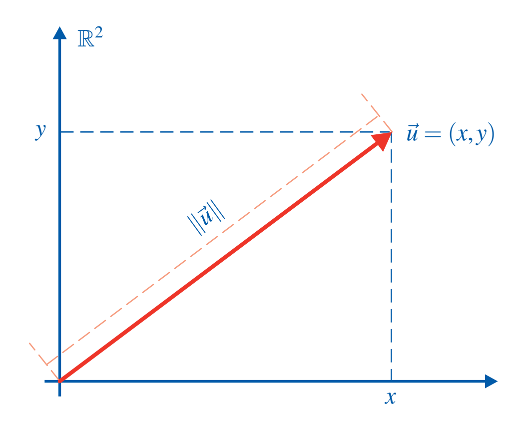


Figura 3-4 • La norma o magnitud de un vector es la longitud del vector.

5. **Distancia entre vectores**. Si $\vec{u}$ y $\vec{v}$ son vectores de $\mathbb{R}^2$, se define la distancia entre ellos como la magnitud del vector $\vec{v} - \vec{u}$; esto es,

$$\mathrm{d}(\vec{v}, \vec{u}) = \|\vec{v} - \vec{u}\|.$$

Note que el vector $\vec{v} - \vec{u}$ es el vector que sumado a $\vec{u}$ da como resultado el vector $\vec{v}$ tal como se ilustra en la figura 3-5, y que la distancia de $\vec{u}$ a $\vec{v}$ es la misma que la distancia de $\vec{v}$ a $\vec{u}$; es decir, $\mathrm{d}(\vec{u}, \vec{v}) = \mathrm{d}(\vec{v}, \vec{u})$.

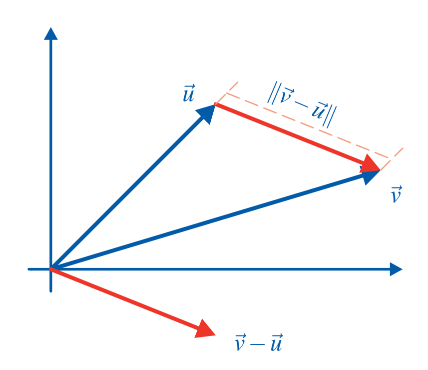


Figura 3-5 • La distancia entre vectores es la magnitud de la diferencia entre ellos.

6. **Producto interior o escalar**. El producto interior o escalar (o producto punto) de $\vec{u}$ con $\vec{v}$, como el lector recordará de sus cursos de física, está dado por

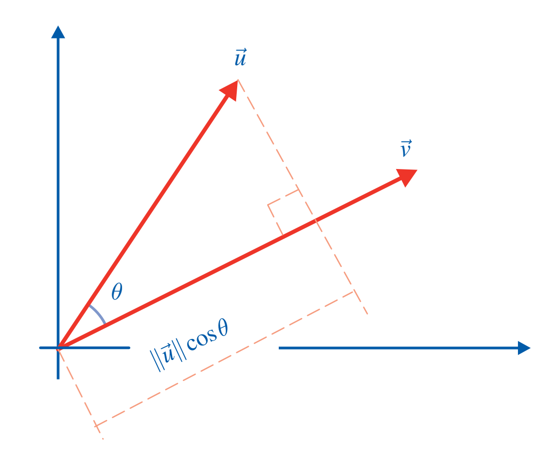
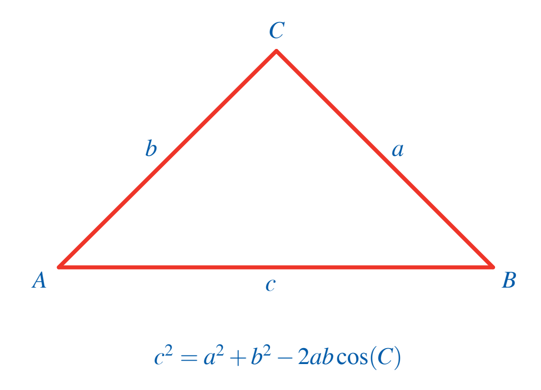


Figura 3-6 • El producto punto de dos vectores es la magnitud de la proyección del primer vector sobre el segundo multiplicada por la norma de este último.

$$\vec{u}\cdot\vec{v} = \|\vec{u}\|\,\|\vec{v}\|\cos\theta \tag{3.1}$$

donde θ es el ángulo entre $\vec{u}$ y $\vec{v}$; y es la magnitud de la proyección de $\vec{u}$ sobre $\vec{v}$ multiplicada por la norma de $\vec{v}$ (véase la figura 3-6). Mediante el producto escalar también se define el trabajo físico.

Observe que todas las características anteriores del espacio $\mathbb{R}^2$ están perfectamente determinadas algebraicamente por las coordenadas (x,y) de los vectores correspondientes, excepto el producto punto. Nos proponemos dar una fórmula alternativa para calcular el producto punto que no dependa de conocer el ángulo entre los vectores; específicamente, deseamos hallar una relación del producto interior que dependa exclusivamente de las componentes de los vectores. Para ello necesitaremos de la llamada ley de los cosenos, conocida por el lector de sus cursos de trigonometría, que recordamos en la figura 3-7.



$$c^2 = a^2 + b^2 - 2ab\cos(C)$$

Figura 3-7 • Si se representan por letras mayúsculas las magnitudes de los ángulos y por letras minúsculas las magnitudes de los correspondientes lados opuestos a cada ángulo, entonces se cumple la relación $c^2 = a^2 + b^2 - 2ab\cos C$, llamada ley de los cosenos, para cualquier triángulo dado.

Ahora sean $\vec{u} = (x_1, y_1)$ y $\vec{v} = (x_2, y_2)$ un par de vectores en $\mathbb{R}^2$. De la figura 3-8 y la ley de los cosenos tenemos que

$$\|\vec{v} - \vec{u}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\|\vec{u}\|\,\|\vec{v}\|\cos\theta.$$

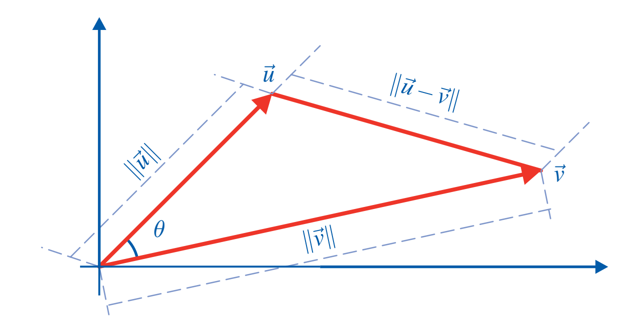


Figura 3-8 •

Luego,

$$\begin{aligned}2\|\vec{u}\|\,\|\vec{v}\|\cos\theta &= \|\vec{u}\|^2+\|\vec{v}\|^2-\|\vec{v}-\vec{u}\|^2\\ &= x_1^2+y_1^2+x_2^2+y_2^2-[(x_2-x_1)^2+(y_2-y_1)^2]\\ &= x_1^2+y_1^2+x_2^2+y_2^2-[x_2^2-2x_1x_2+x_1^2+y_2^2-2y_1y_2+y_1^2].\end{aligned}$$

De donde

$$2\|\vec{u}\|\,\|\vec{v}\|\cos\theta = 2(x_1x_2+y_1y_2)$$

y, por tanto,

$$\vec{u}\cdot\vec{v} = x_1x_2+y_1y_2 \tag{3.2}$$

es la relación buscada.

❍ **Nota 3.1**

1. Observe que $\|\vec{u}\| = \sqrt{x^2+y^2} = \sqrt{\vec{u}\cdot\vec{u}}$.
2. Por otra parte, ya que $-1 \le \cos\theta \le 1$, $\forall\theta\in\mathbb{R}$, se tiene

$$|\vec{u}\cdot\vec{v}| \le \|\vec{u}\|\,\|\vec{v}\|,$$

   a la cual se le llama **desigualdad de Schwarz**.
3. También, de (3.1) (cfr. pág. 116), el **ángulo** entre dos vectores $\vec{u}$, $\vec{v}$ no nulos está dado por:

$$\theta = \arccos\left[\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}\right] \tag{3.3}$$

3.1.2 Interpretación geométrica del determinante

Aunque el determinante es un concepto útil asociado a las matrices, para el estudio de éstas y de otros aspectos del álgebra lineal, el determinante tiene una interpretación geométrica sumamente importante,

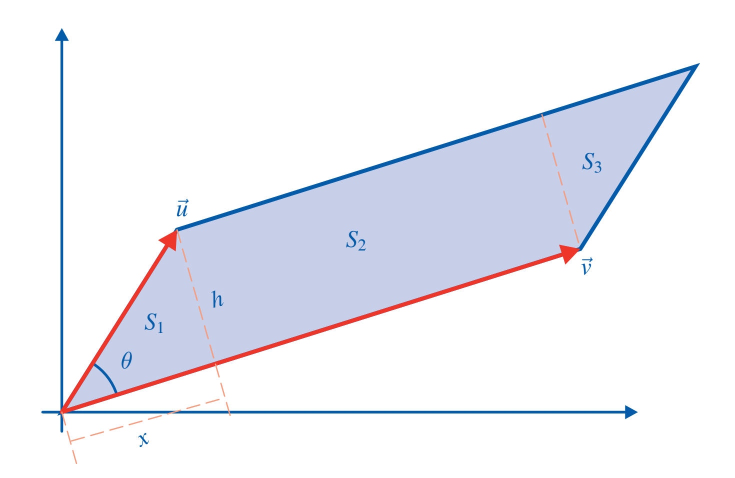


Figura 3-9 •

la cual se puede usar, por ejemplo, en la teoría de integración, específicamente en la fórmula de cambio de variables para integrales y en el cálculo de volúmenes y áreas. Es en estas dos últimas en las cuales enfocaremos nuestra atención.

Sean $\vec{u}=(a,b)$, $\vec{v}=(c,d)$ dos vectores de $\mathbb{R}^2$. Calculemos el área S del paralelogramo generado por estos vectores. Entonces, de acuerdo con la figura 3-9, $S=2S_1+S_2$ (porque $S_1=S_3$). Ahora bien, $S_1=(xh)/2$ y $h=\|\vec{u}\|\operatorname{sen}\theta$. Entonces,

$$\begin{aligned}
S &= 2\frac{xh}{2}+S_2\\
&= xh+S_2\\
&= x\|\vec{u}\|\operatorname{sen}\theta+S_2\\
&= x\|\vec{u}\|\operatorname{sen}\theta+(\|\vec{v}\|-x)h\\
&= x\|\vec{u}\|\operatorname{sen}\theta+h\|\vec{v}\|-xh\\
&= x\|\vec{u}\|\operatorname{sen}\theta+h\|\vec{v}\|-x\|\vec{u}\|\operatorname{sen}\theta\\
&= \|\vec{u}\|\,\|\vec{v}\|\operatorname{sen}\theta.
\end{aligned}$$

Puesto que $\operatorname{sen}^2\theta=1-\cos^2\theta$,

$$\begin{aligned}
S^2 &= \|\vec{u}\|^2\|\vec{v}\|^2\operatorname{sen}^2\theta\\
&= \|\vec{u}\|^2\|\vec{v}\|^2(1-\cos^2\theta)\\
&= \|\vec{u}\|^2\|\vec{v}\|^2-\|\vec{u}\|^2\|\vec{v}\|^2\cos^2\theta\\
&= \|\vec{u}\|^2\|\vec{v}\|^2-(\vec{u}\cdot\vec{v})^2\\
&= \left(a^2+b^2\right)\left(c^2+d^2\right)-(ac+bd)^2\\
&= a^2c^2+a^2d^2+b^2c^2+b^2d^2-a^2c^2-2acbd-b^2d^2\\
&= a^2d^2+b^2c^2-2acbd\\
&= (ad-bc)^2,
\end{aligned}$$

de donde

$$S=|ad-bc|\,.$$

Figura 3-10 • El volumen del paralelepípedo generado por los vectores $\vec{u}$, $\vec{v}$ y $\vec{w}$ es el valor absoluto del determinante de la matriz que tiene como filas (o columnas) a estos vectores.

Es decir,

$$S = |\det(M)|,$$

donde M es la matriz que tiene como filas (o columnas) a los vectores $\vec{u}$ y $\vec{v}$. De manera análoga, el determinante de una matriz 3×3 (o mejor dicho, su valor absoluto) será el volumen del paralelepípedo determinado por los vectores fila de la propia matriz, como se ilustra en la figura 3-10.

3.1.3 El espacio vectorial $\mathbb{R}^n$, geometría y propiedades algebraicas

En esta subsección nos proponemos generalizar las propiedades algebraicas y geométricas de $\mathbb{R}^2$ a espacios de mayor "dimensión". Es evidente que, tanto en la teoría como en la práctica, surgen problemas que involucran un número de variables mayor a dos (en algunos problemas de importancia netamente aplicada, este número puede ser de hasta 20 000). Por tanto, es necesario estudiar aquellos conjuntos cuyos elementos son n-adas ordenadas y tienen cualidades análogas a las de los vectores del plano de coordenadas.

Definición 3.2 *Sea n un número entero positivo. Se define el espacio $\mathbb{R}^n$ como*

$$\mathbb{R}^n = \{\vec{u} = (x_1, x_2, \ldots, x_n) \,|\, x_1, x_2, \ldots, x_n \in \mathbb{R}\}.$$

A los elementos de $\mathbb{R}^n$ también les llamaremos vectores. Al número x_i se le dice la i-ésima ***componente*** *o* ***coordenada*** *de $\vec{u}$. Y al vector $\vec{u}$ lo denotaremos, indistintamente, como $\vec{u} = (x_1, x_2, \ldots, x_n)$ o como la matriz columna*

$$\vec{u} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

❍ **Nota 3.2** Cuando se emplea la notación $\vec{u} = (x_1, x_2, \ldots, x_n)$, se acostumbra decir que $(x_1, x_2, \ldots, x_n)$ es una n-ada ordenada; por ejemplo, (x_1, x_2, x_3) es una tríada ordenada.

Definición 3.3 *Sean $\vec{u}, \vec{v} \in \mathbb{R}^n$, con $\vec{u} = (x_1, x_2, \ldots, x_n)$, $\vec{v} = (y_1, y_2, \ldots, y_n)$ y $\lambda \in \mathbb{R}$.*

1. ***Igualdad.*** *$\vec{u} = \vec{v} \Leftrightarrow x_i = y_i$ para cada $i = 1, 2, \ldots, n$.*
2. ***Suma de vectores.*** *Se denota y define la suma de $\vec{u}$ con $\vec{v}$ como*

$$\vec{u} + \vec{v} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

3. ***Producto de un escalar por un vector.*** *Se denota y define el producto del escalar λ con el vector $\vec{u}$ como*

$$\lambda\vec{u} = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n).$$

4. ***Producto escalar.*** *El producto escalar (o producto punto o producto interior) de $\vec{u}$ con $\vec{v}$ se denota y define como*

$$\vec{u} \cdot \vec{v} = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \cdots + x_n y_n .$$

5. ***Norma de un vector.*** *Se define la norma o magnitud del vector $\vec{u}$ por*

$$\|\vec{u}\| = \left[\sum_{i=1}^{n} x_i^2\right]^{1/2} = \left[x_1^2 + \cdots + x_n^2\right]^{1/2}.$$

6. ***Distancia entre vectores (puntos) en $\mathbb{R}^n$.*** *Se define la distancia entre los vectores $\vec{u}$ y $\vec{v}$ como*

$$\mathrm{d}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| .$$

❍ **Nota 3.3**

1. La relación entre el producto interior y la norma vuelve a ser, como en el caso de $\mathbb{R}^2$,

$$\|\vec{u}\| = (\vec{u} \cdot \vec{u})^{1/2} . \tag{3.4}$$

2. Existen otras formas de medir magnitudes de vectores; sin embargo, la históricamente más común es la que hemos usado hasta ahora dada por (3.4). A esta magnitud se le acostumbra llamar **norma euclidiana** o **norma canónica** del vector $\vec{u}$ por su origen geométrico y natural, respectivamente.
3. Dado que los vectores en $\mathbb{R}^n$ los hemos denotado también como matrices columna, el producto punto se puede ver como producto de matrices; esto es,

$$\vec{u} \cdot \vec{v} = (\vec{u})^t \vec{v}.$$

En el lado izquierdo de la precedente igualdad los vectores están escritos con la notación de n-adas ordenadas y en el lado derecho están escritos con la notación de matrices columna.

Así, vemos que todas las definiciones dadas arriba son generalizaciones directas de la manera en que se opera, se mide y se hace geometría y álgebra en el plano de coordenadas $\mathbb{R}^2$. También debe notarse la importancia que tuvo dar una fórmula alternativa para el producto interior que dependiera sólo de las componentes de los vectores (cfr. fórmula (3.2), pág. 117); de no ser así, no podríamos haber generalizado el producto punto al no tener una forma de "medir ángulos" por medios físicos en estos espacios cuando $n > 3$. Sin embargo, el concepto de ángulo entre vectores sí lo podremos extender a los espacios $\mathbb{R}^n$ con $n > 3$ mediante la desigualdad de Schwarz que veremos más adelante.

▶ **Ejemplo 3.1** Si $n = 3$, $\mathbb{R}^3$ es el espacio usual de tres dimensiones donde "habitamos". En este espacio necesitamos de tres números reales (x_1, x_2, x_3) [o, como tradicionalmente se escribe, (x, y, z)] para determinar la posición de un punto, como hacemos patente en la figura 3-11.◀

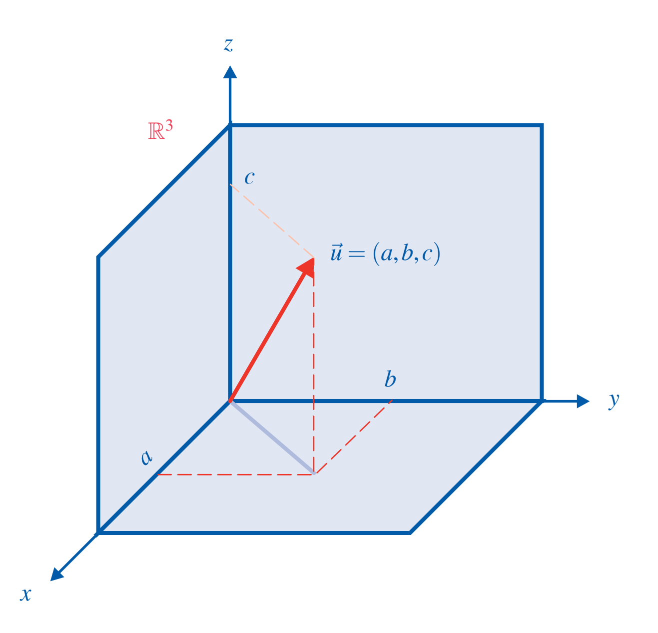


Figura 3-11 • En el espacio $\mathbb{R}^3$ todo punto (vector) $\vec{u}$ se localiza mediante una tríada ordenada (a, b, c); donde las dos primeras componentes (a, b) son la proyección vertical de este punto sobre el plano x, y y la tercera, c, es la proyección horizontal de este punto sobre el eje z.

❍ **Nota 3.4** Es común en textos de matemáticas convenir que cuando se hacen diagramas del espacio tridimensional $\mathbb{R}^3$, los ejes x, y y z se coloquen como en la figura 3-11. Usted puede recordar esta convención (algunas veces llamada *regla de la mano izquierda*) colocando su mano izquierda con la palma frente a usted, abriendo los dedos medio, índice y pulgar (cerrando los dedos anular y meñique); apuntando el dedo medio hacia usted, el índice hacia su derecha y el pulgar hacia arriba. Así, sus dedos señalarán las direcciones positivas del eje x (dedo medio), eje y (dedo índice) y eje z (dedo pulgar).

▶ **Ejemplo 3.2** Si $\vec{u} = (1, -2, 4)$ y $\vec{v} = (3, 6, 0)$, entonces $\vec{u}, \vec{v} \in \mathbb{R}^3$ y:

1. $\vec{u} + \vec{v} = (4, 4, 4)$.
2. $-\sqrt{2}\vec{u} = (-\sqrt{2}, 2\sqrt{2}, -4\sqrt{2})$.
3. $\vec{u} \cdot \vec{v} = (1)(3) + (-2)(6) + (4)(0) = 3 - 12 + 0 = -9$.
4. $\|\vec{u}\| = \sqrt{1^2 + (-2)^2 + 4^2} = \sqrt{21}$.◀

▶ **Ejemplo 3.3** Si $\vec{u} = (-1,2,5,9)$, $\vec{v} = (2,-4,0,3) \in \mathbb{R}^4$:

1. $\vec{u} \cdot \vec{v} = -2 - 8 + 27 = 17$.
2. $\vec{u} + (-2)\vec{v} = (-1,2,5,9) + (-4,8,0,-6) = (-5,10,5,3)$.
3. $\|\vec{v}\| = (4+16+9)^{1/2} = \sqrt{29}$.◀

▶ **Ejemplo 3.4** Calcular la distancia entre los vectores:

1. $\vec{u} = (1,-2,1)$ y $\vec{v} = (-2,1,1)$.
2. $\vec{u} = (-1,0,2,3,1)$ y $\vec{v} = (1,1,-4,2,0)$.◀

Solución 1.
$$\begin{aligned} d(\vec{u},\vec{v}) &= \|\vec{u} - \vec{v}\| \\ &= \|(1,-2,1) - (-2,1,1)\| \\ &= \|(3,-3,0)\| \\ &= 3\sqrt{2}. \end{aligned}$$

2.
$$\begin{aligned} d(\vec{u},\vec{v}) &= \|\vec{u} - \vec{v}\| \\ &= \|(-1,0,2,3,1) - (1,1,-4,2,0)\| \\ &= \|(-2,-1,6,1,1)\| \\ &= \sqrt{43}. \end{aligned}$$ ✔

A continuación enunciamos, sin demostrar, las propiedades algebraicas esenciales de $\mathbb{R}^n$. Utilizando la conmutatividad, asociatividad, etc. de los números reales y la definición de igualdad de vectores, el lector puede fácilmente verificarlas. De hecho, como veremos más adelante, dichas propiedades también se pueden observar en otros conjuntos[1] que llamaremos espacios vectoriales.

Propiedades del espacio vectorial de $\mathbb{R}^n$

Si $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ y λ, $\beta \in \mathbb{R}$, entonces:

1. $\vec{u} + \vec{v} \in \mathbb{R}^n$. (La suma es cerrada)
2. $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$. (Asociatividad de la suma)
3. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$. (Conmutatividad de la suma)
4. Si $\vec{0}_{\mathbb{R}^n} = (\underbrace{0,0,\ldots,0}_{n})$, $\vec{0}_{\mathbb{R}^n} \in \mathbb{R}^n$ y $\vec{u} + \vec{0}_{\mathbb{R}^n} = \vec{u}$, $\forall \vec{u} \in \mathbb{R}^n$. (Existencia del neutro aditivo)
5. Dado $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, existe $-\vec{u} \in \mathbb{R}^n$ tal que $\vec{u} + (-\vec{u}) = \vec{0}_{\mathbb{R}^n}$.
   De hecho, $-\vec{u} = (-x_1, -x_2, \ldots, -x_n)$. (Existencia del inverso aditivo)
6. $\lambda\vec{u} \in \mathbb{R}^n$. (El producto por un escalar es cerrado)
7. $\lambda(\beta\vec{u}) = (\lambda\beta)\vec{u}$. (Asociatividad del producto con escalares)

[1] Compare con las propiedades 1*(a)* a 1*(j)* de la subsección 1.1.4, páginas 9 y 10.

8. $(\lambda+\beta)\vec{u}=\lambda\vec{u}+\beta\vec{u}$. (Distributividad del producto con respecto a la suma de escalares)
9. $\lambda(\vec{u}+\vec{v})=\lambda\vec{u}+\lambda\vec{v}$. (Distributividad del producto con respecto a la suma de vectores)
10. $1\vec{u}=\vec{u}$, $\forall\vec{u}\in\mathbb{R}^n$. (Preservación de la escala)

El producto escalar tiene las siguientes importantes propiedades que son sencillas de probar y cuya demostración se deja al lector.

Propiedades del producto punto

Sean $\vec{u},\vec{v},\vec{w}\in\mathbb{R}^n$ y $\lambda\in\mathbb{R}$, entonces:

1. $\vec{u}\cdot\vec{v}=\vec{v}\cdot\vec{u}$. (Simetría)
2. $\vec{u}\cdot(\lambda\vec{v})=\lambda(\vec{u}\cdot\vec{v})$. (Homogeneidad)
3. $\vec{u}\cdot(\vec{v}+\vec{w})=\vec{u}\cdot\vec{v}+\vec{u}\cdot\vec{w}$. (Distributividad)
4. $\vec{u}\cdot\vec{u}\geq 0$ y $\vec{u}\cdot\vec{u}=0\Leftrightarrow\vec{u}=\vec{0}$. (Positividad)

3.1.4 La desigualdad de Schwarz, ángulos entre vectores y ortogonalidad

Es claro que la generalización natural de ángulo entre vectores en $\mathbb{R}^n$ debe estar dada por la fórmula[2]

$$\theta=\arccos\left[\frac{\vec{u}\cdot\vec{v}}{\|u\|\,\|v\|}\right].$$

Pero, para que se pueda evaluar la función arccos, es necesario que

$$-1\leq\frac{\vec{u}\cdot\vec{v}}{\|u\|\,\|v\|}\leq 1$$

lo cual evidentemente equivale a

$$|\vec{u}\cdot\vec{v}|\leq\|\vec{u}\|\,\|\vec{v}\|.$$

Afortunadamente, esta desigualdad es cierta y la probaremos en el teorema 3.1. Estaremos entonces facultados para generalizar el concepto de ángulo entre vectores en $\mathbb{R}^n$ para $n>3$. Para poder demostrar dicha desigualdad necesitamos de la sencilla proposición que damos a continuación (lema 3.1).

Lema 3.1 *Si a, b son cualquier par de números reales, entonces*

$$2ab\leq a^2+b^2. \tag{3.5}$$

[2] Cfr. (3.3), página 117.

DEMOSTRACIÓN ■ En efecto,

$$\begin{aligned} 0 &\le (a-b)^2 \\ &= a^2 - 2ab + b^2; \end{aligned}$$

de donde $2ab \le a^2 + b^2$. ■

Teorema 3.1 ***(Desigualdad de Schwarz)*** *Si $\vec{u} = (x_1, x_2, \ldots, x_n)$, $\vec{v} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, entonces*

$$|\vec{u} \cdot \vec{v}| \le \|\vec{u}\| \, \|\vec{v}\| \tag{3.6}$$

DEMOSTRACIÓN ■ Si $\vec{u} = \vec{0}$ o $\vec{v} = \vec{0}$, claramente (3.6) es cierta. Supongamos que $\vec{u}, \vec{v} \neq \vec{0}$; sea $i \in \{1, 2, \ldots, n\}$ arbitrario y pongamos

$$a = \frac{x_i}{\|\vec{u}\|} \text{ y } b = \frac{y_i}{\|\vec{v}\|}$$

en el lema 3.1. Entonces, por (3.5), para todo $i = 1, 2, \ldots, n$ se tiene

$$2 \frac{x_i}{\|\vec{u}\|} \frac{y_i}{\|\vec{v}\|} \le \left(\frac{x_i}{\|\vec{u}\|}\right)^2 + \left(\frac{y_i}{\|\vec{v}\|}\right)^2 .$$

Luego,

$$2 \sum_{i=1}^{n} \frac{x_i}{\|\vec{u}\|} \frac{y_i}{\|\vec{v}\|} \le \sum_{i=1}^{n} \left[\left(\frac{x_i}{\|\vec{u}\|}\right)^2 + \left(\frac{y_i}{\|\vec{v}\|}\right)^2 \right];$$

esto es,

$$\frac{2}{\|\vec{u}\| \, \|\vec{v}\|} \sum_{i=1}^{n} x_i y_i \le \sum_{i=1}^{n} \left(\frac{x_i}{\|\vec{u}\|}\right)^2 + \sum_{i=1}^{n} \left(\frac{x_i}{\|\vec{u}\|}\right)^2 ,$$

que equivale a

$$\begin{aligned} \frac{2}{\|\vec{u}\| \, \|\vec{v}\|} \sum_{i=1}^{n} x_i y_i &\le \frac{1}{\|\vec{u}\|^2} \sum_{i=1}^{n} x_i^2 + \frac{1}{\|\vec{v}\|^2} \sum_{i=1}^{n} y_i^2 \\ &= \frac{1}{\|\vec{u}\|^2} \|\vec{u}\|^2 + \frac{1}{\|\vec{v}\|^2} \|\vec{v}\|^2 \\ &= 2 \end{aligned}$$

de donde

$$\sum_{i=1}^{n} x_i y_i \le \|\vec{u}\| \, \|\vec{v}\| .$$

Con lo que hemos probado

$$\vec{u} \cdot \vec{v} \le \|\vec{u}\| \, \|\vec{v}\| \qquad \forall \vec{u}, \vec{v} \tag{3.7}$$

Entonces, por la homogeneidad del producto punto y la precedente desigualdad, tenemos

$$-(\vec{u} \cdot \vec{v}) = (-\vec{u}) \cdot \vec{v} \le \|-\vec{u}\| \, \|\vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| ;$$

que significa

$$\vec{u}\cdot\vec{v} \geq -\,\|\vec{u}\|\,\|\vec{v}\| \tag{3.8}$$

De (3.7) y (3.8) se deduce

$$|\vec{u}\cdot\vec{v}| \leq \|\vec{u}\|\,\|\vec{v}\| \qquad \forall \vec{u},\vec{v}\in\mathbb{R}^n\,. \quad \blacksquare$$

Ahora sí podemos definir, con base en la desigualdad de Schwarz (3.6), el ángulo entre vectores de n componentes; que es una generalización del concepto de ángulo entre vectores en el espacio de tres dimensiones y en el plano cartesiano.

Definición 3.4 *Si $\vec{u}$, $\vec{v}\in\mathbb{R}^n-\{\vec{0}_{\mathbb{R}^n}\}$, se define el* **ángulo** *entre estos vectores como:*

$$\theta = \arccos\left[\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}\right] \tag{3.9}$$

► **Ejemplo 3.5** Hallar el ángulo θ entre los vectores $(1,2,0,2)$ y $(-3,1,1,5)$ de $\mathbb{R}^4$.◄

Solución

$$\begin{aligned}\theta &= \arccos\left[\frac{(1,2,0,2)\cdot(-3,1,1,5)}{\sqrt{1^2+2^2+0^2+2^2}\sqrt{(-3)^2+1^2+1^2+5^2}}\right]\\ &= \arccos\left[\frac{9}{3\cdot 6}\right] = \arccos\left[\frac{1}{2}\right].\end{aligned}$$

Así, $\theta = 60^\circ$. ✔

Una vez que se ha definido el concepto de ángulo entre vectores, se puede determinar cuándo un par de éstos son perpendiculares; la manera de generalizar esta idea a $\mathbb{R}^n$ la hacemos patente a continuación. Notemos, de (3.9), que el ángulo entre dos vectores es de 90° si y sólo si su producto punto es cero.

Definición 3.5

1. *Dos vectores $\vec{u}$, $\vec{v}\in\mathbb{R}^n$ son* **ortogonales** *(perpendiculares) si*

$$\vec{u}\cdot\vec{v} = 0\,;$$

   es decir, si el ángulo entre ellos es de 90°. Cuando $\vec{u}$ y $\vec{v}$ sean ortogonales lo denotaremos por $\vec{u}\perp\vec{v}$.

2. *$\vec{u}$ y $\vec{v}$ son paralelos ($\vec{u}\,\|\,\vec{v}$) si $\vec{u}\cdot\vec{v} = \pm\|\vec{u}\|\,\|\vec{v}\|$, lo que equivale a que el ángulo entre ellos sea de 0° o 180°.*

Teorema de Pitágoras

Quizá uno de los más importantes y conspicuos resultados de las matemáticas, conocido y usado en forma empírica desde el inicio de la civilización (Babilonia, Egipto), luego convertido en una afirmación general y probado en forma completamente rigurosa por los griegos, es el teorema de Pitágoras.

Se enseña en educación elemental y, a partir de ahí, se hace uso sistemático de él. Sin embargo, la mayor parte de la gente que usa este teorema desconoce alguna demostración porque se requieren varios resultados elementales de geometría para poder establecer una prueba rigurosa. A continuación daremos una demostración muy simple de este importante teorema; de hecho es trivial, pues el material que hemos desarrollado hasta aquí nos proporciona una herramienta algebraica muy potente para atacar este problema geométrico transformándolo en un sencillo cálculo algebraico.

Teorema 3.2 ***(Teorema de Pitágoras)*** *Sean $\vec{u}, \vec{v} \in \mathbb{R}^n$ un par de vectores ortogonales. Entonces,*

$$\|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 .$$

Antes de dar la demostración de este teorema explicaremos su relación con el teorema de Pitágoras que el lector conoce, pues en apariencia no hay una relación directa. Sin embargo, si particularizamos al caso $n = 2$ del plano cartesiano $\mathbb{R}^2$, resulta que $\|\vec{u}+\vec{v}\|$ es la longitud de la hipotenusa del triángulo rectángulo, cuyos catetos tienen longitudes $\|\vec{u}\|$ y $\|\vec{v}\|$, como se ilustra en la figura 3-12.

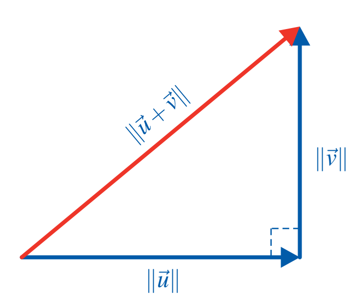


Figura 3-12 • El teorema de Pitágoras establece que en cualquier triángulo rectángulo la suma de los cuadrados de los catetos es igual al cuadrado de la hipotenusa.

DEMOSTRACIÓN ■ Dado que $\vec{u} \perp \vec{v}$, se tiene $\vec{u} \cdot \vec{v} = 0$; luego

$$\begin{aligned} \|\vec{u}+\vec{v}\|^2 &= (\vec{u}+\vec{v}) \cdot (\vec{u}+\vec{v}) \\ &= \vec{u} \cdot \vec{u} + \vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{u} + \vec{v} \cdot \vec{v} \\ &= \|\vec{u}\|^2 + 2\vec{u} \cdot \vec{v} + \|\vec{v}\|^2 \\ &= \|\vec{u}\|^2 + 0 + \|\vec{v}\|^2 \\ &= \|\vec{u}\|^2 + \|\vec{v}\|^2 . \quad \blacksquare \end{aligned}$$

Propiedades de la norma en $\mathbb{R}^n$

Uno de los conceptos más importantes en matemáticas es el de proximidad; y la forma de medir la proximidad entre puntos es por medio de la distancia. En $\mathbb{R}$, el valor absoluto es la herramienta usada para medir la distancia entre números reales. Recordemos que el valor absoluto tiene las siguientes propiedades:

1. $|x| \geq 0 \; \forall x \in \mathbb{R}$.
2. $|x| = 0 \Leftrightarrow x = 0$.

3. $|\lambda x| = |\lambda|\,|x|$.
4. $|x+y| \le |x| + |y|$. (Desigualdad triangular)

Entonces la distancia entre un par de números x y y se define como el valor absoluto de su diferencia. En $\mathbb{R}^n$, la norma es la manera natural para definir proximidad entre vectores por medio de la distancia entre ellos. A partir de las propiedades 1, 2, 3 y 4 del valor absoluto, enunciadas arriba, se pueden deducir todas las demás propiedades que tiene el valor absoluto. De hecho, estas mismas propiedades las tiene la norma en $\mathbb{R}^n$ y cualquier otra propiedad de la norma también se puede deducir a partir de éstas.

Teorema 3.3 ***(Propiedades de la norma en*** $\mathbb{R}^n$***)*** *La norma en* $\mathbb{R}^n$ *tiene las siguientes propiedades:*

1. $\|\vec{u}\| \ge 0 \quad \forall \vec{u} \in \mathbb{R}^n$.
2. $\|\vec{u}\| = 0 \Leftrightarrow \vec{u} = \vec{0}_{\mathbb{R}^n}$.
3. $\|\lambda\vec{u}\| = |\lambda|\,\|\vec{u}\| \quad \forall \vec{u} \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$.
4. $\|\vec{u}+\vec{v}\| \le \|\vec{u}\| + \|\vec{v}\| \quad \forall \vec{u}, \vec{v} \in \mathbb{R}^n$. *(*Desigualdad triangular*)*

DEMOSTRACIÓN ■ 1. Si $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, entonces

$$\|\vec{u}\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \ge 0\,.$$

2. Si $\vec{u} = \vec{0}_{\mathbb{R}^n}$, claramente $\|\vec{u}\| = 0$. Supongamos inversamente que $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ es tal que $\|\vec{u}\| = 0$, entonces

$$\sum_{i=1}^{n} x_i^2 = 0;$$

de donde $x_i^2 = 0\ \forall i = 1, 2, \ldots, n$; y por tanto $\vec{u} = \vec{0}_{\mathbb{R}^n}$.

3. Si $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$, entonces

$$\begin{aligned}
\|\lambda\vec{u}\| &= \|(\lambda x_1, \lambda x_2, \ldots, \lambda x_n)\| \\
&= \left(\sum_{i=1}^{n} (\lambda x_i)^2\right)^{1/2} \\
&= \left(\sum_{i=1}^{n} \lambda^2 x_i^2\right)^{1/2} \\
&= \left(\lambda^2 \sum_{i=1}^{n} x_i^2\right)^{1/2} \\
&= |\lambda| \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \\
&= |\lambda|\,\|\vec{u}\|\,.
\end{aligned}$$

4. Si $\vec{u},\vec{v}\in\mathbb{R}^n$, entonces

$$\begin{aligned}\|\vec{u}+\vec{v}\|^2 &= (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v})\\ &= \vec{u}\cdot\vec{u}+\vec{u}\cdot\vec{v}+\vec{v}\cdot\vec{u}+\vec{v}\cdot\vec{v}\\ &= \|\vec{u}\|^2+2(\vec{u}\cdot\vec{v})+\|\vec{v}\|^2.\end{aligned}$$

De la desigualdad de Schwarz se tiene que

$$\vec{u}\cdot\vec{v}\le\|\vec{u}\|\,\|\vec{v}\|$$

y por tanto

$$\|\vec{u}+\vec{v}\|^2\le\|\vec{u}\|^2+2\|\vec{u}\|\,\|\vec{v}\|+\|\vec{v}\|^2;$$

esto es,

$$\|\vec{u}+\vec{v}\|^2\le(\|\vec{u}\|+\|\vec{v}\|)^2.$$

De donde,

$$\|\vec{u}+\vec{v}\|\le\|\vec{u}\|+\|\vec{v}\|.\quad\blacksquare$$

La desigualdad triangular recibe este nombre porque, en el caso de vectores en $\mathbb{R}^2$ (o en $\mathbb{R}^3$), significa que *en todo triángulo, la longitud de cualquiera de sus lados es inferior a la suma de las longitudes de los otros dos lados*; como ilustramos en la figura 3-13.

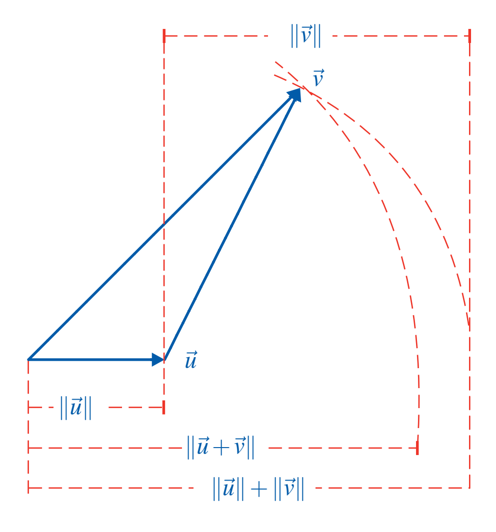


Figura 3-13 • Desigualdad triangular: en todo triángulo, la longitud de cualquiera de sus lados es inferior a la suma de las longitudes de sus otros dos lados.

Planos en $\mathbb{R}^3$

Supongamos que un plano P pasa por el punto $\vec{u}_0=(x_0,y_0,z_0)$ y es ortogonal al vector $\vec{\eta}=(a,b,c)$; es decir, $\vec{\eta}$ es perpendicular a toda línea recta contenida en el plano P. Sea $\vec{u}=(x,y,z)$ un punto cualquiera

del plano P (cfr. figura 3-14). Entonces $\vec{\eta}$ es perpendicular al segmento que une a los puntos $\vec{u}_0$ y $\vec{u}$; es decir, $\vec{\eta} \perp (\vec{u} - \vec{u}_0)$. Por tanto, $\vec{\eta} \cdot (\vec{u} - \vec{u}_0) = 0$ y, por ende,

$$a\,(x - x_0) + b\,(y - y_0) + c\,(z - z_0) = 0 \tag{3.10}$$

es la ecuación que determina el lugar geométrico correspondiente al plano P. Esto significa que todo punto (x, y, z) que pertenece al plano P satisface la ecuación (3.10). Inversamente, toda solución (x, y, z) de esta ecuación pertenece al plano P.

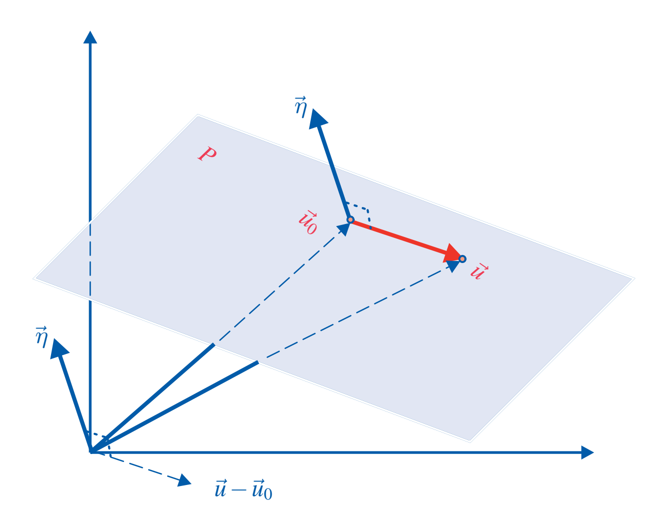


Figura 3-14 • Un plano P que pasa por un punto dado $\vec{u} = (x_0, y_0, z_0)$ y es ortogonal a un vector $\vec{\eta} = (a, b, c)$.

Es claro que la ecuación (3.10) equivale[3] a

$$ax + by + cz = d \tag{3.11}$$

donde $d = ax_0 + by_0 + cz_0$.

► **Ejemplo 3.6** Encontrar la ecuación del plano que es ortogonal al vector $\vec{\eta} = (-1, 2, 4)$ y pasa por el punto $\vec{u} = (2, 1, 1)$. ◄

Solución La ecuación está dada por (3.10):

$$(-1)\,(x - 2) + 2\,(y - 1) + 4\,(z - 1) = 0$$

que equivale a

$$-x + 2y + 4z = 4. \quad ✔$$

[3] Es obvio que las ecuaciones (3.10) y (3.11) no son únicas en cuanto a la descripción algebraica de un plano como lugar geométrico. En realidad, cualquier otra ecuación algebraica que cumpla con ese objetivo será equivalente a (3.10) y (3.11) en el sentido de que tienen las mismas soluciones y, por ende, describen el mismo lugar geométrico: el plano en cuestión.

▶ **Ejemplo 3.7** Encontrar la ecuación del plano P que pasa por los puntos $\vec{u} = (1,0,0)$, $\vec{v} = (0,1,0)$ y $\vec{w} = (0,0,1)$.◀

Solución El plano debe contener los tres puntos y, por la definición geométrica de plano, también debe contener los segmentos de línea que unen a dichos puntos; dos de ellos son $\overline{\vec{u}\vec{v}}$ y $\overline{\vec{u}\vec{w}}$. Si $\vec{\eta} = (a,b,c)$ es un vector ortogonal al plano, entonces $\vec{\eta}$ debe ser perpendicular a estos dos segmentos. Así que $\vec{\eta} \perp (\vec{u} - \vec{v})$ y $\vec{\eta} \perp (\vec{u} - \vec{w})$, por lo que $\vec{\eta} \cdot (\vec{u} - \vec{v}) = 0$ y $\vec{\eta} \cdot (\vec{u} - \vec{w}) = 0$. Esto es,

$$\begin{aligned} (1,-1,0) \cdot (a,b,c) &= 0 \\ (1,0,-1) \cdot (a,b,c) &= 0; \end{aligned}$$

es decir,

$$\begin{aligned} a - b &= 0 \\ a - c &= 0. \end{aligned}$$

Resolvamos ahora el sistema homogéneo anterior:

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix},$$

que produce las soluciones

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} r \\ r \\ r \end{bmatrix}; r \in \mathbb{R}.$$

Una solución particular es el vector $\vec{\eta} = (1,1,1)$ obtenida al hacer $r = 1$. Así, el plano que pasa por estos puntos es ortogonal al vector $\vec{\eta} = (1,1,1)$, y contiene al punto $(1,0,0)$; por tanto, al utilizar (3.10), tenemos

$$(1)(x-1) + (1)(y-0) + (1)(z-0) = 0,$$

que equivale a

$$x + y + z = 1.$$

Este plano viene bosquejado en la figura 3-15. ✔

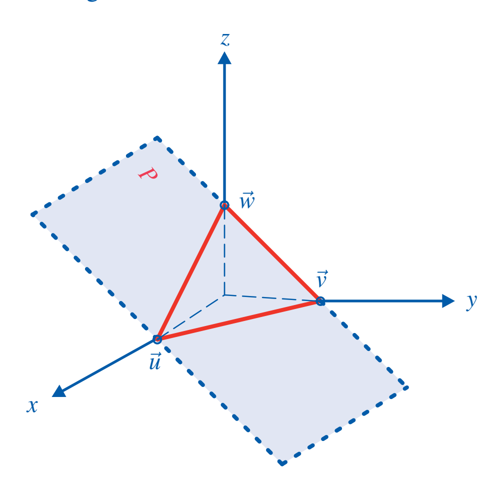


Figura 3-15 • Plano que pasa por los puntos $\vec{u} = (1,0,0)$, $\vec{v} = (0,1,0)$ y $\vec{w} = (0,0,1)$.

3.2 Espacios vectoriales

Hemos visto que las matrices tienen, con la suma y el producto por un escalar usuales, las mismas diez propiedades[4] que las del espacio vectorial[5] de $\mathbb{R}^n$. Así, las generalizaciones algebraicas del plano y del espacio tienen un símil con un conjunto aparentemente sin conexión con ellos. Como veremos en esta sección, el de las matrices no es un caso aislado y, al contrario, existe una gran variedad de conjuntos en los que se han definido las operaciones suma entre sus elementos y multiplicación de números reales con estos elementos, que también satisfacen las citadas diez condiciones. Todos estos conjuntos tienen en común las propiedades mencionadas y, por tanto, lo que se derive de ellas dependerá de las mismas y no de los elementos que particularmente formen determinada colección. Por ello surge la necesidad de estudiar este tipo de conjuntos, con sus respectivas operaciones, en abstracto y no caso por caso de manera aislada. De esta forma, lo que haremos primeramente es abstraer estas diez propiedades como característica esencial de lo que llamaremos *espacio vectorial*, y entonces podremos derivar, a partir de las mismas, consecuencias generales en este espacio abstracto que serán entonces válidas, dado que dependen solamente de las propiedades de las operaciones y no de los elementos de cada colección, en todos los conjuntos que cumplan con esas diez propiedades.

3.2.1 Definiciones y ejemplos

Definición 3.6 *Sea* $\mathbf{E}$ *un conjunto no vacío donde se han definido un par de operaciones: suma entre sus elementos, representada como* $\vec{u}\oplus\vec{v}$*; y multiplicación (o producto) de escalares (números reales) con elementos de* $\mathbf{E}$*, representada como* $\alpha\odot\vec{u}$*. Entonces a* $\mathbf{E}$ *se le llama* **espacio vectorial** (real) *si se cumplen las diez siguientes condiciones (axiomas de espacio vectorial):*

1. $\vec{u}\oplus\vec{v}\in\mathbf{E}\quad\forall\vec{u},\vec{v}\in\mathbf{E}$. (La suma es cerrada)
2. $\vec{u}\oplus(\vec{v}\oplus\vec{w})=(\vec{u}\oplus\vec{v})\oplus\vec{w}\quad\forall\vec{u},\vec{v},\vec{w}\in\mathbf{E}$. (Asociatividad de la suma)
3. $\vec{u}\oplus\vec{v}=\vec{v}\oplus\vec{u}\quad\forall\vec{u},\vec{v}\in\mathbf{E}$. (Conmutatividad de la suma)
4. Existe un elemento $\vec{0}_{\mathbf{E}}\in\mathbf{E}$ tal que $\vec{u}\oplus\vec{0}_{\mathbf{E}}=\vec{u}\quad\forall\vec{u}\in\mathbf{E}$. (Existencia del neutro aditivo)
5. Para cada $\vec{u}\in\mathbf{E}$ existe $-\vec{u}\in\mathbf{E}$ *tal que* $\vec{u}\oplus(-\vec{u})=\vec{0}_{\mathbf{E}}$. (Existencia del inverso aditivo)
6. $\lambda\odot\vec{u}\in\mathbf{E}\quad\forall\lambda\in\mathbb{R},\forall\vec{u}\in\mathbf{E}$. (La multiplicación con escalares es cerrada)
7. $\lambda\odot(\beta\odot\vec{u})=(\lambda\beta)\odot\vec{u}\quad\forall\lambda,\beta\in\mathbb{R},\forall\vec{u}\in\mathbf{E}$. (Asociatividad del producto con escalares)
8. $\lambda\odot(\vec{u}\oplus\vec{v})=(\lambda\odot\vec{u})\oplus(\lambda\odot\vec{v})\quad\forall\lambda\in\mathbb{R},\forall\vec{u},\vec{v}\in\mathbf{E}$. (Distributividad del producto con respecto a la suma de vectores)
9. $(\lambda+\beta)\odot\vec{u}=(\lambda\odot\vec{u})\oplus(\beta\odot\vec{u})\quad\forall\lambda,\beta\in\mathbb{R},\forall\vec{u}\in\mathbf{E}$. (Distributividad del producto con respecto a la suma de escalares)
10. $1\odot\vec{u}=\vec{u}\quad\forall\vec{u}\in\mathbf{E}$. (Preservación de la escala)

A los elementos de $\mathbf{E}$ *les llamaremos vectores.*

[4] Cfr. propiedades 1*(a)* a 1*(j)*, subsección 1.1.4, página 9.
[5] Cfr. subsección 3.1.3, página 119.

► **Ejemplo 3.8** $\mathbb{R}^n$, con la suma usual de vectores y la multiplicación de un escalar por un vector, esto es, si $\vec{u} = (x_1, x_2, \ldots, x_n)$, $\vec{v} = (y_1, y_2, \ldots, y_n)$ y λ es un número real,

$$\begin{aligned} \vec{u} \oplus \vec{v} = (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n), \\ \lambda \odot \vec{u} = \lambda(x_1, x_2, \ldots, x_n) &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n), \end{aligned}$$

es un espacio vectorial.◄

► **Ejemplo 3.9** Sea $\mathbf{E} = \{x \in \mathbb{R} \mid x > 0\}$ dotado de las operaciones:

$$\begin{aligned} u \oplus v &= uv \\ \lambda \odot u &= u^\lambda. \end{aligned}$$

Así, por ejemplo, $2 \oplus 3 = 2 \cdot 3 = 6$ y $2 \odot 3 = 3^2 = 9$.

¿Es $\mathbf{E}$, junto con estas operaciones, un espacio vectorial?◄

Solución Para contestar afirmativamente debemos probar que se verifican las diez propiedades de la definición anterior; y para dar una respuesta negativa se tiene que exhibir un caso en el cual una de ellas (por lo menos) no se cumpla.

1. Es claro de su definición que $u \oplus v \in \mathbf{E}$ $\forall u, v \in \mathbf{E}$, pues el producto de dos números positivos es también positivo.
2. $\forall u, v, w \in \mathbf{E} : (u \oplus v) \oplus w = (uv) \oplus w = (uv)w = u(vw) = u \oplus (vw) = u \oplus (v \oplus w)$.
3. $\forall\, u, v \in \mathbf{E}$: $u \oplus v = uv = vu = v \oplus u$.
4. Sea $\vec{0}_\mathbf{E} = 1 \in \mathbf{E}$. Entonces $\forall\, u \in \mathbf{E}$: $u \oplus \vec{0}_\mathbf{E} = u\vec{0}_\mathbf{E} = u1 = u$.
5. Si $u \in \mathbf{E}$, $u > 0$, y por tanto $1/u > 0$. Sea $-u = 1/u$. Entonces $-u \in \mathbf{E}$ y $u \oplus (-u) = u(1/u) = 1 = \vec{0}_\mathbf{E}$.
6. Si $\lambda \in \mathbb{R}$ y $u \in \mathbf{E}$, entonces $\lambda \odot u = u^\lambda > 0$ pues $u > 0$, es decir, $\lambda \odot u \in \mathbf{E}$ $\forall u$, $\forall\, \lambda \in \mathbb{R}$.
7. Si $\lambda, \beta \in \mathbb{R}$ y $u \in \mathbf{E}$, $\lambda \odot (\beta \odot u) = \lambda \odot (u^\beta) = (u^\beta)^\lambda = u^{\beta\lambda} = (\lambda\beta) \odot u$.
8. $\forall u, v \in \mathbf{E}, \lambda \in \mathbb{R}$: $\lambda \odot (u \oplus v) = \lambda \odot (uv) = (uv)^\lambda = u^\lambda v^\lambda = (\lambda \odot u) \oplus (\lambda \odot v)$.
9. $\forall\, \lambda, \beta \in \mathbb{R}$, $\forall\, u \in \mathbf{E}$: $(\lambda + \beta) \odot u = u^{\lambda+\beta} = u^\lambda u^\beta = (\lambda \odot u) \oplus (\beta \odot u)$.
10. $1 \odot u = u^1 = u$, $\forall u \in \mathbf{E}$.

Como se verifican los diez axiomas de espacio vectorial de la definición 3.6, $\mathbf{E}$ es un espacio vectorial. ✔

► **Ejemplo 3.10** (Espacio de matrices) $\mathfrak{M}_{m \times n}$, el conjunto de matrices tamaño $m \times n$, con la suma de matrices y el producto de un escalar por una matriz usuales, es un espacio vectorial.[6]◄

[6] Cfr. página 9.

▶ **Ejemplo 3.11** (Espacio de polinomios) Si **P** es el conjunto de polinomios, con la suma y el producto por un escalar usuales, **P** es un espacio vectorial; lo cual es fácil de probar y se deja de ejercicio al lector.◀

❍ **Nota 3.5**

1. Observemos que en el ejemplo 3.9 no utilizamos la notación de poner una flecha encima de los vectores del espacio; pues las operaciones, como fueron definidas, involucraron el producto y la multiplicación de números reales positivos y el haber denotado a los elementos de esta manera podría haber causado confusión. En realidad, esto lo haremos cada vez que sea conveniente; así, por ejemplo, las matrices sólo las denotaremos con letras mayúsculas en lugar de emplear notación vectorial para el espacio $\mathfrak{M}_{m\times n}$. También, a partir del final de esta nota, abandonaremos la notaciones $\oplus$ y $\odot$ para la suma de vectores y la multiplicación de un escalar por un vector y simplemente escribiremos $\vec{u}+\vec{v}$, $\lambda\vec{u}$ en lugar de $\vec{u}\oplus\vec{v}$ y $\lambda\odot\vec{u}$, respectivamente; pues el contexto de cada caso evitará cualquier confusión.
2. El axioma 10 de la definición 3.6 es en apariencia una propiedad de la que se podría prescindir; ya que es una característica que *a simple vista* se cumple "siempre" de manera "natural". Sin embargo, esta propiedad es imprescindible; pues existen casos en los que se pueden definir operaciones de suma de vectores y multiplicación con escalares que cumplen con los primeros 9 axiomas de la definición 3.6, pero no con el número 10. Por ejemplo, si en $\mathbb{R}^n$ se define la suma de vectores en forma usual, pero el producto por un escalar como $\lambda\odot\vec{u}=\vec{0}_{\mathbb{R}^n}$, para todo $\lambda\in\mathbb{R}$ y para todo $\vec{u}\in\mathbb{R}^n$, es evidente que se cumplen los 9 primeros axiomas de espacio vectorial pero no el décimo. También es claro que este último conjunto con esas operaciones no tiene trascendencia alguna como son los casos de las matrices y el propio espacio $\mathbb{R}^n$ con las operaciones usuales. Evitar casos triviales y de nulo interés en la práctica es la razón de ser del axioma 10 para el concepto de espacio vectorial.

▶ **Ejemplo 3.12** (Espacio de sucesiones) Sea $\mathbb{R}^\infty=\left\{\vec{u}=(a_n)_{n\in\mathbb{N}}\,|\,\vec{u}\text{ es una sucesión}\right\}$; es decir, $\mathbb{R}^\infty$ es el conjunto de las sucesiones de números reales $\vec{u}=(a_1,a_2,\ldots,a_n,\ldots)$. Con la suma de sucesiones y el producto de un escalar por una sucesión usuales; esto es,

- $(a_1,a_2,\ldots,a_n,\ldots)+(b_1,b_2,\ldots,b_n,\ldots)=(a_1+b_1,a_2+b_2,\ldots,a_n+b_n,\ldots)$,
- $\lambda(a_1,a_2,\ldots,a_n,\ldots)=(\lambda a_1,\lambda a_2,\ldots,\lambda a_n,\ldots)$.

$\mathbb{R}^\infty$ es un espacio vectorial como fácilmente puede comprobar el lector.◀

Espacios de funciones

Definición 3.7 *Sean A y B un par de conjuntos no vacíos. Una función f con dominio A y valores en B, es una regla que a cada elemento x de A le asigna un único elemento $y=f(x)$ de B. Para denotar que f es una función con dominio A y valores en B escribiremos*

$$f:A\longrightarrow B.$$

La función o regla de asignación es f y no debe confundirse con el valor de ésta en x: $y=f(x)$. En esta última notación, a y se le dice la variable dependiente o imagen de x bajo f, x es la variable independiente o argumento de la función f, y a B se le denomina contradominio de la función.

Definición 3.8 *Sean $f,g : A \longrightarrow B$ un par de funciones de A en B. Diremos que $f = g$, si*

$$f(x) = g(x) \qquad \forall x \in A.$$

► **Ejemplo 3.13** Sean $A = \mathbb{R} - \{0\}$, $B = \mathbb{R}$ y $f,g : A \longrightarrow B$ las funciones definidas como

$$f(x) = \frac{|x|}{x}$$

y

$$g(x) = \begin{cases} 1, & \text{si } x > 0; \\ -1, & \text{si } x < 0. \end{cases}$$

Si $x \in A$ y $x > 0$,

$$f(x) = \frac{|x|}{x} = \frac{x}{x} = 1 = g(x).$$

Si $x \in A$ y $x < 0$,

$$f(x) = \frac{|x|}{x} = \frac{-x}{x} = -1 = g(x).$$

Luego $f(x) = g(x)$ $\forall x \in A$, por tanto $f = g$.◄

Definición 3.9 *Sea A un conjunto no vacío. Denotamos por $\mathscr{F}(A)$ al conjunto de las funciones $f : A \longrightarrow \mathbb{R}$. Dotamos a $\mathscr{F}(A)$ de las siguientes operaciones:*

- *Si $f,g \in \mathscr{F}(A)$, definimos la función $f+g$ como $(f+g)(x) = f(x)+g(x)$ $\forall x \in A$.*
- *Si $f \in \mathscr{F}(A)$ y $\lambda \in \mathbb{R}$, definimos la función λf como $(\lambda f)(x) = \lambda f(x)$ $\forall x \in A$.*

Diremos, entonces, que éstas son la suma y la multiplicación por un escalar usuales en las funciones con dominio A y valores reales.

► **Ejemplo 3.14** Sea $A = \{x \,|\, x$ es elemento del grupo MA0084302$\}$, y $f,g \in \mathscr{F}(A)$ las funciones $f(x)$ = matrícula de x; $g(x)$ = calificación de x en el primer examen parcial de la materia de álgebra lineal. Entonces, si la matrícula de Liliana es 447021, Liliana obtuvo 10 en el primer examen parcial, x = Liliana y $\lambda = .5$:

$$(f+g)(x) = f(x)+g(x) = 447021+10 = 447031,$$

$$(\lambda g)(x) = \lambda g(\text{Liliana}) = .5 \cdot 10 = 5.$$

Es decir, la función $f+g$ es la regla que asigna la suma de la matrícula y la calificación del primer parcial a cada elemento de la clase de álgebra lineal. Mientras que la función λg asigna a cada elemento de A el producto de λ con su calificación del primer parcial.◄

❍ **Nota 3.6** Sean a, b un par de números reales, $a < b$ y $A = [a,b]$, y sea $f : A \to \mathbb{R}$ una función. La letra f, en este caso, representa la regla de asociación entre los elementos x de A y los valores asignados $y = f(x)$. Una ventaja que tienen las funciones de este tipo es que se pueden identificar con un ente geométrico que es la gráfica de dicha función. En la figura 3-16 se bosqueja (hipotéticamente) la gráfica de la función f, ella consiste en los pares ordenados $(x, f(x))$ con $x \in [a,b]$; que al ubicarlos en el plano cuando x recorre el intervalo $[a,b]$ forman la curva bosquejada en esta figura. Es conveniente que el lector tenga siempre presente, cuando trabaje con funciones de este tipo, identificar la función con su gráfica para fines de tener ideas concretas de la misma y no confundir el valor de la función f en x, que hemos representado como $y = f(x)$, con la propia función f.

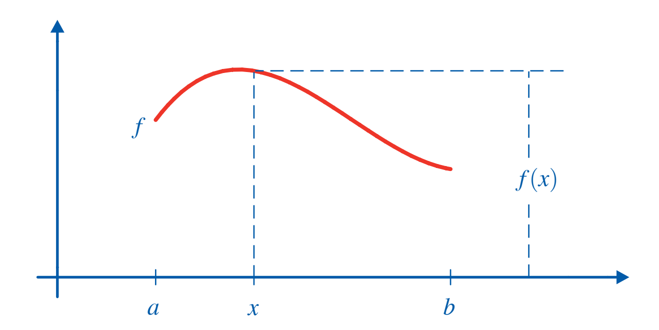


Figura 3-16 • La gráfica de una función se puede identificar con la propia función.

Sean a, b números reales, con $a < b$ y $A = [a,b]$. Sean $f, g : A \to \mathbb{R}$ un par de funciones. Entonces la función $f + g$ evaluada en $x \in A$, es decir, $(f+g)(x)$, se obtiene sumando los valores de f en x y g en x. Esto es,[7] $(f+g)(x) = f(x) + g(x)$. La gráfica de $f+g$ se obtiene entonces sumando las "alturas" $f(x)$ y $g(x)$ en cada $x \in A$ y bosquejando el punto $(x, f(x)+g(x))$ como se muestra en la figura 3-17. Así podemos pensar, geométricamente, que la función $f+g$ es la correspondiente curva mostrada en esta figura.

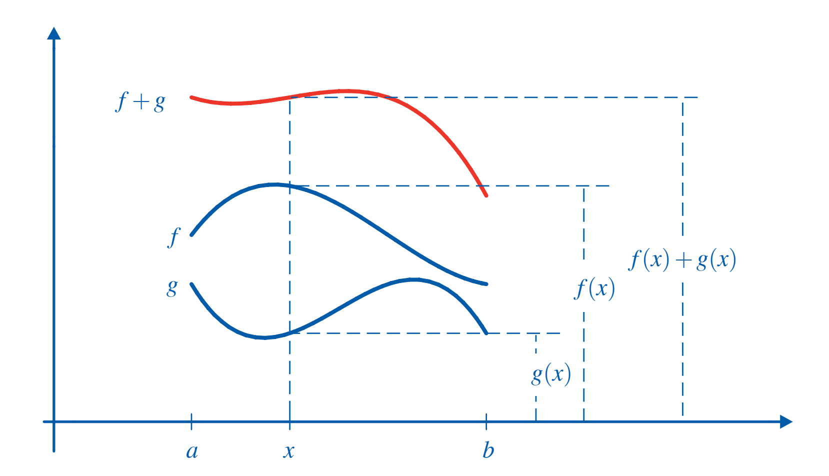


Figura 3-17 • Las funciones f, g y $f+g$ representadas por sus respectivas gráficas.

[7] El lector debe tener mucho cuidado en no confundirse al pensar que $(f+g)(x) = f(x) + g(x)$ es "una distribución de un producto con la suma de números"; pues esta interpretación evidentemente no tiene sentido.

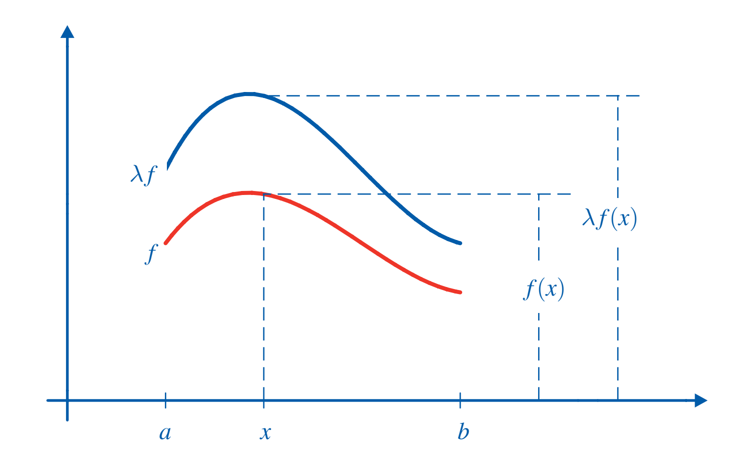


Figura 3-18 • Las funciones f y λf representadas por sus correspondientes gráficas.

De manera análoga, si λ es un número real y $f : A \to \mathbb{R}$, la función λf se evalúa, en cada $x \in A$, mediante el producto de números reales $\lambda f(x)$. De esta forma la gráfica de la función λf se obtiene, a partir de la gráfica de la función f, multiplicando cada una de las "alturas" $f(x)$ por λ, en cada $x \in A$, y bosquejando los puntos $(x, \lambda f(x))$ como se ilustra en la figura 3-18.

► **Ejemplo 3.15** (Espacio de funciones) Si A es cualquier conjunto no vacío mostrar que $\mathscr{F}(A)$, junto con las operaciones dadas en la definición 3.9, es un espacio vectorial; el llamado *espacio de las funciones con dominio A y valores reales*.◄

DEMOSTRACIÓN ■ 1. Claramente, de la definición de suma de funciones:

$$f, g \in \mathscr{F}(A) \Rightarrow f + g \in \mathscr{F}(A).$$

2. Si $f, g, h \in \mathscr{F}(A)$ y x es cualquier elemento de A,

$$[f + (g + h)](x) = f(x) + (g + h)(x) = f(x) + (g(x) + h(x)) = (f(x) + g(x)) + h(x),$$

así que,[8] $f + (g + h) = (f + g) + h$.

3. Si $f, g \in \mathscr{F}(A)$ y $x \in A$ es cualquier elemento, tenemos

$$(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x),$$

y, por ende, $f + g = g + f$.

4. Sea $\theta : A \longrightarrow \mathbb{R}$ la función en $\mathscr{F}(A)$, definida como $\theta(x) = 0\ \forall x \in A$. Entonces, para toda $f \in \mathscr{F}(A)$ y para todo $x \in A$:

$$(f + \theta)(x) = f(x) + \theta(x) = f(x) + 0 = f(x),$$

por ende, $f + \theta = f$.

[8] Aquí hemos utilizado los hechos de que $f(x)$, $g(x)$, $h(x)$ son números reales y en éstos hay asociatividad. El lector debe notar que en el transcurso de esta demostración utilizaremos las propiedades de los números reales y que los valores de las funciones son también reales.

5. Si $f \in \mathscr{F}(A)$, sea $-f : A \longrightarrow \mathbb{R}$ la función definida, para cada $x \in A$, como $(-f)(x) = -f(x)$. Entonces

$$(f+(-f))(x) = f(x)+(-f)(x) = f(x)+(-f(x)) = f(x)-f(x) = 0 = \theta(x),$$

$\forall x \in A$; es decir, $f+(-f) = \theta$.

6. Claramente, $\lambda f \in \mathscr{F}(A)\ \forall f \in \mathscr{F}(A)$ y $\forall \lambda \in \mathbb{R}$.
7. Si $\lambda, \beta \in \mathbb{R}$ y $f \in \mathscr{F}(A)$, para cada $x \in A$,

$$(\lambda(\beta f))(x) = \lambda(\beta f)(x) = \lambda(\beta f(x)) = (\lambda\beta)f(x),$$

de donde $\lambda(\beta f) = (\lambda\beta)f$.

8. $\forall \lambda \in \mathbb{R}, \forall f, g \in \mathscr{F}(A)$ y para cada $x \in A$:

$$[\lambda(f+g)](x) = \lambda(f+g)(x) = \lambda(f(x)+g(x)) = \lambda f(x)+\lambda g(x) = (\lambda f+\lambda g)(x);$$

es decir, $\lambda(f+g) = \lambda f+\lambda g$.

9. Si $\lambda, \beta \in \mathbb{R}$ y $f \in \mathscr{F}(A)$, para todo $x \in A$,

$$((\lambda+\beta)f)(x) = (\lambda+\beta)f(x) = \lambda f(x)+\beta f(x) = (\lambda f+\beta f)(x),$$

lo que implica $(\lambda+\beta)f = \lambda f+\beta f$.

10. Si $f \in \mathscr{F}(A)$, para todo $x \in A$,

$$(1f)(x) = 1f(x) = f(x),$$

es decir, $1f = f$. ■

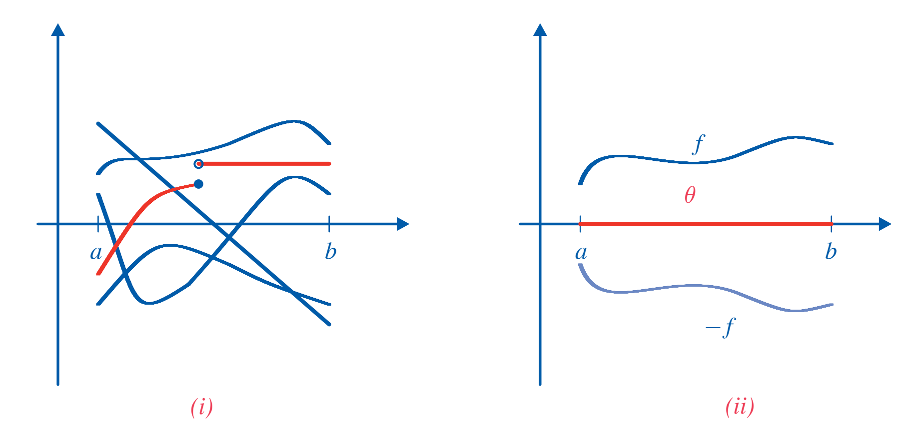


Figura 3-19 • *(i)* El espacio vectorial $\mathscr{F}([a,b])$ consta de todas las funciones cuya gráfica está contenida en la franja (x,y) con $x \in [a,b]$ y $y \in \mathbb{R}$, es decir, la franja $[a,b] \times \mathbb{R}$. *(ii)* El neutro aditivo de este espacio, θ, es la función constante cero ($\theta(x) = 0\ \forall x \in [a,b]$), cuya gráfica se encuentra en color rojo; y el inverso aditivo de una función f se encuentra reflejando su gráfica sobre el eje x.

Cuando $A = [a,b]$ es un intervalo, $\mathscr{F}(A)$ es el conjunto de todas las funciones cuya gráfica está contenida en la franja $[a,b] \times \mathbb{R}$. En la figura 3-19*(i)* hemos bosquejado algunos de sus elementos. En la

misma figura 3-19*(ii)* ilustramos la interpretación geométrica del neutro aditivo θ, la función constante cero (en rojo) del espacio $\mathscr{F}([a,b])$. También en esta última figura hemos bosquejado el inverso aditivo para una función dada f; el cual se obtiene gráficamente reflejando sobre el eje x la gráfica de la función f, como el lector fácilmente puede comprobar por sí mismo recordando la interpretación dada en la figura 3-17 para la suma usual de funciones.

❍ **Nota 3.7** Cabe hacer notar que la demostración de que $\mathscr{F}(A)$ sea un espacio vectorial no depende de las características del conjunto A, sino del hecho de que sus elementos son funciones con valores en $\mathbb{R}$ y de que los números reales tienen estructura de espacio vectorial. Por ende, si sustituimos $\mathbb{R}$ por un espacio vectorial arbitrario, $\mathscr{F}(A)$ seguirá siendo un espacio vectorial. Lo mismo sucede si ponemos cualquier conjunto concreto en lugar de A.

3.2.2 Propiedades elementales de los espacios vectoriales

Como ya mencionamos al principio de esta sección, trabajaremos con un espacio vectorial abstracto para establecer propiedades que se infieran únicamente de los diez axiomas de espacio vectorial dados en la definición 3.6, y entonces estas deducciones serán válidas para cualquier espacio vectorial concreto. Llevaremos esto a cabo a través de lo que resta de este libro y no volveremos a hacer más énfasis en esta estrategia natural. Las primeras propiedades quedan establecidas en el siguiente teorema.

Teorema 3.4 ***(Propiedades elementales de espacio vectorial)*** *Sea* $\mathbf{E}$ *un espacio vectorial. Entonces:*

1. $\vec{0}_{\mathbf{E}}$ *es único. Es decir, sólo existe un elemento en* $\mathbf{E}$ *que al sumarlo con cualquier otro elemento del espacio da como resultado este último.*
2. $-\vec{u}$ *es único para cada* $\vec{u}\in\mathbf{E}$*. Esto es, cada elemento del espacio tiene exactamente un inverso aditivo.*
3. $\vec{u}+\vec{v}=\vec{u}+\vec{w}\Rightarrow\vec{v}=\vec{w}$*. A esta propiedad se le llama* **ley de cancelación**.
4. $0\vec{u}=\vec{0}_{\mathbf{E}}$ *para todo* $\vec{u}\in\mathbf{E}$.
5. $\lambda\vec{0}_{\mathbf{E}}=\vec{0}_{\mathbf{E}} \quad \forall\lambda\in\mathbb{R}$.
6. $(-1)\vec{u}=-\vec{u} \quad \forall\vec{u}\in\mathbf{E}$.
7. $\vec{u}+\vec{u}=2\vec{u} \quad \forall\vec{u}\in\mathbf{E}$.
8. *Si* n *es un entero,* $\underbrace{\vec{u}+\vec{u}+\cdots+\vec{u}}_{n}=n\vec{u} \quad \forall\vec{u}\in\mathbf{E}$.
9. *Si* $\alpha\in\mathbb{R}$ *y* $\vec{u}\in\mathbf{E}$*, entonces*

$$\alpha\vec{u}=\vec{0}_{\mathbf{E}}\Leftrightarrow\alpha=0 \quad \text{o} \quad \vec{u}=\vec{0}_{\mathbf{E}}.$$

DEMOSTRACIÓN ■ 1. Sea $\varphi\in\mathbf{E}$ tal que $\vec{u}+\varphi=\vec{u}\;\forall\,\vec{u}\in\mathbf{E}$, entoces $\varphi=\vec{0}_{\mathbf{E}}+\varphi=\vec{0}_{\mathbf{E}}$; por tanto, $\varphi=\vec{0}_{\mathbf{E}}$.

2. Si $\vec{u}\in\mathbf{E}$ y $\vec{u}_1\in\mathbf{E}$ satisface $\vec{u}+\vec{u}_1=\vec{0}_{\mathbf{E}}$, entonces

$$\vec{u}_1=\vec{u}_1+\vec{0}_{\mathbf{E}}=\vec{u}_1+(\vec{u}+(-\vec{u}))=(\vec{u}_1+\vec{u})+(-\vec{u})=\vec{0}_{\mathbf{E}}+(-\vec{u})=-\vec{u}.$$

3.
$$\begin{aligned} \vec{u}+\vec{v} &= \vec{u}+\vec{w} \Rightarrow \\ ((-\vec{u})+\vec{u})+\vec{v} &= ((-\vec{u})+\vec{u})+\vec{w} \Rightarrow \\ \vec{0}_{\mathbf{E}}+\vec{v} &= \vec{0}_{\mathbf{E}}+\vec{w} \Rightarrow \\ \vec{v} &= \vec{w}. \end{aligned}$$

4. $0\vec{u} = (0+0)\vec{u} = 0\vec{u}+0\vec{u}$, y ya que $0\vec{u} = 0\vec{u}+\vec{0}_{\mathbf{E}}$, tenemos
$$\vec{0}_{\mathbf{E}}+0\vec{u} = 0\vec{u}+0\vec{u};$$
de la propiedad (3) se tiene $\vec{0}_{\mathbf{E}} = 0\vec{u}$.
5. $\lambda\vec{0}_{\mathbf{E}} = \lambda(0\vec{0}_{\mathbf{E}}) = (\lambda\cdot 0)\vec{0}_{\mathbf{E}} = 0\vec{0}_{\mathbf{E}} = \vec{0}_{\mathbf{E}}$.
6. $\vec{u}+(-1)\vec{u} = (1+(-1))\vec{u} = 0\vec{u} = \vec{0}_{\mathbf{E}}$, y dado que $-\vec{u}$ es único, $-1\,\vec{u} = -\vec{u}$.
7. $\vec{u}+\vec{u} = (1+1)\vec{u} = 2\vec{u}$.
8. Análoga a la anterior.
9. Se deja de ejercicio al lector. ■

3.2.3 Subespacios vectoriales

Aunque hemos visto varios ejemplos de espacios vectoriales, éstos han sido muy generales y, por lo mismo, demasiado "grandes" para ser útiles. En realidad, estaremos interesados en subconjuntos particulares de estos espacios que, junto con las mismas operaciones, sean también espacios vectoriales, a los cuales llamaremos subespacios vectoriales y a su estudio está dedicado este apartado.

Definición 3.10 *Sean* $\mathbf{E}$ *un espacio vectorial y* S *un subconjunto no vacío de* $\mathbf{E}$*.* S *es un* **subespacio vectorial** *de* $\mathbf{E}$*, o simplemente un subespacio de* $\mathbf{E}$*, si* S*, con las mismas operaciones suma y producto por un escalar restringidas a los elementos de* S*, es también un espacio vectorial. Usaremos la notación* $S < \mathbf{E}$ *para indicar que* S *es un subespacio de* $\mathbf{E}$*.*

Para probar que $S < \mathbf{E}$, se debe mostrar que los elementos de S con la suma y el producto por un escalar definidos en $\mathbf{E}$ satisfacen los diez axiomas de la definición 3.6 de espacio vectorial; sin embargo, si reflexionamos un poco, vemos que las propiedades de asociatividad para la suma y el producto por un escalar, las de distributividad, conmutatividad y preservación de la escala son propiedades que se heredan de $\mathbf{E}$; esto es, puesto que se cumplen para todos los elementos de $\mathbf{E}$ y $S \subset \mathbf{E}$, también son válidas para todos los vectores de S. Luego, resta sólo probar que la suma y el producto por un escalar son operaciones cerradas en S, $\vec{0}_{\mathbf{E}} \in S$ y que para cada $\vec{u} \in S$, $-\vec{u}$ también pertenece a S. Pero si la multiplicación por escalares es cerrada en S, $\vec{u} \in S \Rightarrow -\vec{u} = (-1)\vec{u} \in S$; también, dado que $S \neq \emptyset$, y $0\vec{u} = \vec{0}_{\mathbf{E}}$, se tiene que $\vec{0}_{\mathbf{E}} \in S$. Con esto hemos demostrado el siguiente teorema.

Teorema 3.5 *Sea* S *un subconjunto de un espacio vectorial* $\mathbf{E}$*. Para que* S *sea un subespacio vectorial de* $\mathbf{E}$ *es necesario y suficiente que se cumplan las siguientes tres condiciones:*

1. $\vec{0}_{\mathbf{E}} \in S$.
2. $\vec{u}+\vec{v} \in S\ \forall \vec{u},\vec{v} \in S$.
3. $\lambda\vec{u} \in S\ \forall\lambda \in \mathbb{R},\ \forall\vec{u} \in S$.

▶ **Ejemplo 3.16** Es claro que si **E** es un espacio vectorial, entonces **E** y $\{\vec{0}_{\mathbf{E}}\}$ son subespacios de **E**.◀

❍ **Nota 3.8** Es costumbre llamar a **E** y $\{\vec{0}_{\mathbf{E}}\}$ los subespacios triviales del espacio vectorial **E**; y si $S < \mathbf{E}$ no es **E**, se dice que S es un subespacio propio de **E**.

▶ **Ejemplo 3.17** Sea $S = \{\vec{u} \in \mathbb{R}^2 \,|\, \vec{u} = (x, 2x); x \in \mathbb{R}\}$. Entonces:

1. $\vec{0}_{\mathbb{R}^2} = (0,0) = (0, 2\cdot 0) \in S$.
2. Si $\vec{u}, \vec{v} \in S$, $\vec{u} = (x, 2x)$, $\vec{v} = (y, 2y)$; así que

$$\vec{u} + \vec{v} = (x,2x) + (y,2y) = (x+y, 2x+2y) = (x+y, 2(x+y)) \in S.$$

3. Si $\lambda \in \mathbb{R}$ y $\vec{v} = (x, 2x) \in S, \lambda\vec{u} = \lambda(x, 2x) = (\lambda x, \lambda 2x) = (\lambda x, 2(\lambda x)) \in S$.

De 1, 2 y 3, $S < \mathbb{R}^2$. Geométricamente, S es la línea recta con pendiente 2 que pasa por el origen.◀

De hecho, toda línea recta que pasa por el origen es un subespacio de $\mathbb{R}^2$. Inversamente no es difícil probar, lo cual hará el lector en los ejercicios, que todo subespacio propio de $\mathbb{R}^2$ es geométricamente una línea recta que pasa por el origen o el subespacio trivial $\{(0,0)\}$.

▶ **Ejemplo 3.18** Cualquier plano que contiene al origen y las líneas rectas que pasan por el mismo son subespacios de $\mathbb{R}^3$ (véase la figura 3-20). Más aún, los únicos subespacios no triviales de $\mathbb{R}^3$ son estos conjuntos. (Demuéstrelo.)◀

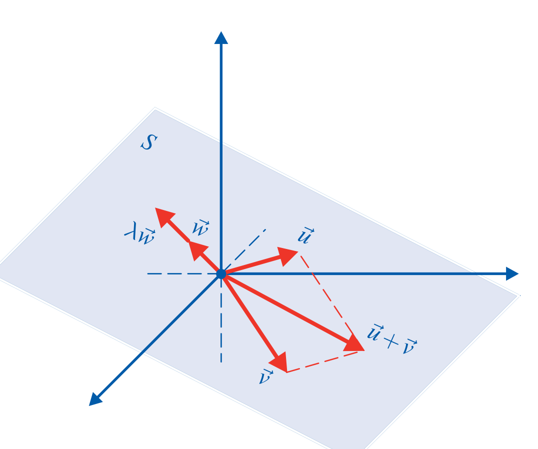


Figura 3-20 • En un plano S que pasa por el origen, la suma de los vectores contenidos en él pertenece a este plano y la multiplicación de los vectores de este plano por escalares sigue perteneciendo a dicho plano. Así, un plano que pasa por el origen es un subespacio de $\mathbb{R}^3$.

▶ **Ejemplo 3.19** Sean a, b, c números reales dados y $S = \{(x,y,z) \in \mathbb{R}^3 | ax+by+cz = 0\}$. Es decir, S es un plano que pasa por el origen con ecuación $ax+by+cz = 0$.

1. Claramente $(0,0,0) \in S$.
2. Si $(x_1,y_1,z_1),(x_2,y_2,z_2) \in S$, entonces $ax_1+by_1+cz_1 = 0$, $ax_2+by_2+cz_2 = 0$ y

$$(x_1,y_1,z_1)+(x_2,y_2,z_2) = (x_1+x_2, y_1+y_2, z_1+z_2);$$

luego,

$$\begin{aligned} a(x_1+x_2)+b(y_1+y_2)+c(z_1+z_2) &= ax_1+by_1+cz_1+ax_2+by_2+cz_2 \\ &= 0+0=0. \end{aligned}$$

Por tanto, $(x_1,y_1,z_1)+(x_2,y_2,z_2)\in S$.

3. Si $\lambda\in\mathbb{R}$, y $\vec{u}=(x_1,y_1,z_1)\in S$, $\lambda\vec{u}=(\lambda x_1,\lambda y_1,\lambda z_1)$; entonces

$$\begin{aligned} a(\lambda x_1)+b(\lambda y_1)+c(\lambda z_1) &= \lambda(ax_1+by_1+z_1) \\ &= \lambda\cdot 0=0. \end{aligned}$$

Por ende, $\lambda\vec{u}\in S$.

De 1, 2 y 3, $S<\mathbb{R}^3$.◄

❍ **Nota 3.9** Observe que un plano S que no pasa por el origen no es un subespacio de $\mathbb{R}^3$, como se hace patente en la figura 3-21.

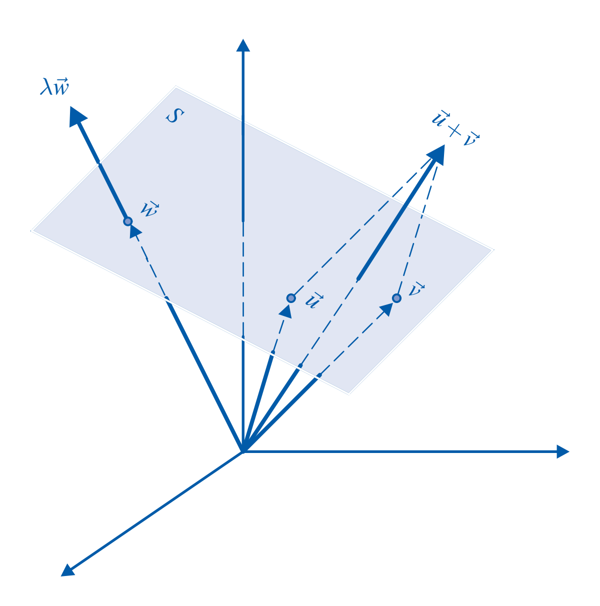


Figura 3-21 • En un plano S que no pasa por el origen $\vec{0}_{\mathbb{R}^n}\notin S$, la suma de vectores contenidos en él no pertenece a este plano y la multiplicación de vectores de este plano por escalares tampoco pertenece necesariamente a dicho plano. Así, un plano que no pasa por el origen no es un subespacio de $\mathbb{R}^3$.

► **Ejemplo 3.20** (Espacio solución, espacio nulo) Si $A\in\mathfrak{M}_{m\times n}$, el conjunto, S, de vectores $\vec{x}$ en $\mathbb{R}^n$ que es solución del sistema homogéneo $A\vec{x}=\vec{0}_{\mathbb{R}^m}$ es un subespacio de $\mathbb{R}^n$, llamado el **subespacio solución** de dicho sistema o el **espacio nulo** de la matriz A.◄

DEMOSTRACIÓN ■ 1. $A\vec{0}_{\mathbb{R}^n}=\vec{0}_{\mathbb{R}^m}$, por tanto, $\vec{0}_{\mathbb{R}^n}\in S$.

2. Si $\vec{x}$, $\vec{y}\in S$, $A\vec{x}=\vec{0}_{\mathbb{R}^m}$ y $A\vec{y}=\vec{0}_{\mathbb{R}^m}$, por tanto, $A(\vec{x}+\vec{y})=A\vec{x}+A\vec{y}=\vec{0}_{\mathbb{R}^m}$. Así, $\vec{x}+\vec{y}\in S$.

3. Si $\lambda\in\mathbb{R}$ y $\vec{x}\in S$, $A\vec{x}=\vec{0}_{\mathbb{R}^m}$. Entonces $A(\lambda\vec{x})=\lambda(A\vec{x})=\lambda\vec{0}_{\mathbb{R}^m}=\vec{0}_{\mathbb{R}^m}$, por ende, $\lambda\vec{x}\in S$. ■

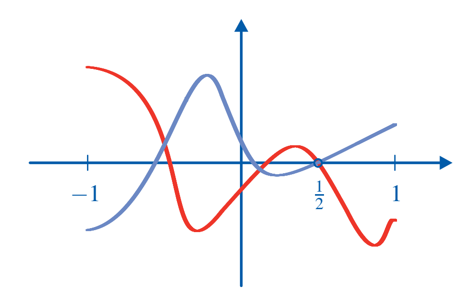


Figura 3-22 • Este espacio consta de todas las funciones cuyas gráficas pasan por el punto $(1/2,0)$ y están contenidas en la franja $[-1,1]\times(-\infty,\infty)$.

► **Ejemplo 3.21** (Espacio de polinomios de grado menor o igual a k). Sea $\mathbf{P}$ el espacio de polinomios del ejemplo 3.11 (pág. 133), y sean k un entero positivo y

$$\mathbf{P}_k = \{p \in \mathbf{P} \,|\, p \text{ tiene grado} \leq k \text{ o es el polinomio constante cero}\}.$$

Entonces, ya que por su definición contiene al cero de $\mathbf{P}$, la suma de dos polinomios de grado a lo más k tiene grado menor o igual a k, y el producto de un polinomio en este conjunto por un escalar tiene grado igual a k o es el polinomio cero, $\mathbf{P}_k < \mathbf{P}$.◄

► **Ejemplo 3.22** Si $\mathbf{E} = \{f : [-1,1] \to \mathbb{R} \,|\, f(1/2) = 0\}$, mostrar que, con la suma y el producto por un escalar usuales en las funciones, $\mathbf{E}$ es un espacio vectorial (véase la figura 3-22).◄

DEMOSTRACIÓN ■ En efecto, ya que $\mathbf{E} \subset \mathscr{F}([-1,1])$ y éste es un espacio vectorial con las operaciones usuales, basta probar que $\mathbf{E} < \mathscr{F}([-1,1])$.

1. $\theta : [-1,1] \to \mathbb{R}$, $\theta(x) = 0\ \forall x \in [-1,1]$, en particular $\theta(1/2) = 0$; por tanto $\theta \in S$.
2. Si $f,g \in \mathbf{E}$, $f(1/2) = 0$, $g(1/2) = 0$; luego $(f+g)(1/2) = f(1/2) + g(1/2) = 0 + 0 = 0$. Así que $f + g \in \mathbf{E}$.
3. Si $\lambda \in \mathbb{R}$ y $f \in \mathbf{E}$, entonces $f(1/2) = 0$ y $(\lambda f)(1/2) = \lambda f(1/2) = \lambda \cdot 0 = 0$; por tanto $\lambda f \in \mathbf{E}$.

De 1, 2 y 3, $\mathbf{E} < \mathscr{F}([-1,1])$. ■

► **Ejemplo 3.23** (Espacio de funciones continuas). Dado que la función cero es una función continua, la suma de funciones continuas y el producto de un escalar por una función continua dan como resultado también funciones continuas, se tiene que

$$C[a,b] = \{f : [a,b] \longrightarrow \mathbb{R} \,|\, f \text{ es continua}\}$$

es un subespacio vectorial de $\mathscr{F}([a,b])$; y por tanto, un espacio vectorial.◄

► **Ejemplo 3.24** De manera análoga, el lector puede probar que[9]

$$C^1[a,b] = \{f : [a,b] \longrightarrow \mathbb{R} \,|\, f \text{ es derivable con continuidad en } [a,b]\}$$

es un espacio vectorial con las operaciones usuales en el espacio de funciones.◄

[9] La continuidad y la derivabilidad en los extremos a y b se sobreentiende que es continuidad y derivabilidad por la derecha y por la izquierda, respectivamente.

▶ **Ejemplo 3.25** (Espacio de matrices simétricas). Sea $\mathfrak{M}_n$ el espacio vectorial de las matrices cuadradas de orden n con las operaciones usuales. Y sea S el subconjunto de matrices simétricas de tamaño $n \times n$. Demostrar que S es un subespacio de $\mathfrak{M}_n$.◀

DEMOSTRACIÓN ■ 1. Claramente la matriz cero de orden n es simétrica.

2. Sean $A, B \in S$, entonces $A^t = A$ y $B^t = B$. Luego, por la tercera propiedad de las matrices (cfr. pág. 10),

$$(A+B)^t = A^t + B^t = A + B;$$

por lo que $A + B \in S$.

3. Si $A \in S$ y $\lambda \in \mathbb{R}$, claramente $(\lambda A)^t = \lambda A^t$. Por tanto, ya que $A \in S$,

$$(\lambda A)^t = \lambda A^t = \lambda A;$$

por lo que $\lambda A \in S$.

De 1, 2 y 3 se concluye que S es un subespacio vectorial de $\mathfrak{M}_n$. ■

3.2.4 Combinaciones lineales y subespacios generados

Supongamos que tenemos un par de vectores no colineales $\vec{u}$ y $\vec{v}$ en $\mathbb{R}^3$ y se multiplican, respectivamente, por los escalares λ y β. El vector resultante, $\lambda\vec{u} + \beta\vec{v}$, está en el plano que pasa por el origen y contiene a estos dos vectores; es decir, el plano que contiene a $\vec{0}$, $\vec{u}$, $\vec{v}$, el cual es un subespacio de $\mathbb{R}^3$ pues es un plano que pasa por el origen (cfr. ejemplos 3.18 y 3.19 pág. 140). Generalizamos a continuación esta idea.

Definición 3.11 *Si* $\mathbf{E}$ *es un espacio vectorial y* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k \in \mathbf{E}$, *a todo vector de la forma* $\vec{v} = a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_k\vec{u}_k$, *donde* $a_1, a_2, \ldots, a_k \in \mathbb{R}$, *se le llama* ***combinación lineal*** *de* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$.

▶ **Ejemplo 3.26** Determinar si el vector $\vec{u} = (7, 1, 16)$ es combinación lineal de los vectores $\vec{u}_1 = (-1, 2, 2)$ y $\vec{u}_2 = (3, -1, 4)$.◀

Solución Veamos si es posible encontrar $a_1, a_2 \in \mathbb{R}$ tales que $\vec{u} = a_1\vec{u}_1 + a_2\vec{u}_2$; esto es,

$$\begin{aligned}(7,1,16) &= a_1(-1,2,2) + a_2(3,-1,4)\\ &= (-a_1 + 3a_2, 2a_1 - a_2, 2a_1 + 4a_2).\end{aligned}$$

Entonces, se debe tener

$$\begin{aligned}-a_1 + 3a_2 &= 7,\\ 2a_1 - a_2 &= 1,\\ 2a_1 + 4a_2 &= 16.\end{aligned}$$

Resolvamos este sistema:

$$\left[\begin{array}{rr|r} -1 & 3 & 7 \\ 2 & -1 & 1 \\ 2 & 4 & 16 \end{array}\right] \sim \left[\begin{array}{rr|r} -1 & 3 & 7 \\ 0 & 5 & 15 \\ 0 & 10 & 30 \end{array}\right]$$
$$\sim \left[\begin{array}{rr|r} -1 & 3 & 7 \\ 0 & 5 & 15 \\ 0 & 0 & 0 \end{array}\right];$$

por tanto,

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Con lo que $\vec{u}$ sí es combinación lineal de $\vec{u}_1$ y $\vec{u}_2$:

$$\vec{u} = 2\vec{u}_1 + 3\vec{u}_2.$$

Comprobación:

$$\begin{aligned} 2\vec{u}_1 + 3\vec{u}_2 &= 2(-1,2,2) + 3(3,-1,4) \\ &= (7,1,16) \\ &= \vec{u}. \quad ✔ \end{aligned}$$

Teorema 3.6 ***(Subespacios generados)*** *Sean* $\mathbf{E}$ *un espacio vectorial y* $\vec{u}_1$, $\vec{u}_2$,..., $\vec{u}_k$ *vectores de* $\mathbf{E}$. *Entonces*

$$S = \{\vec{u} \in \mathbf{E} \,|\, \vec{u} \text{ es combinación lineal de } \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$$

es un subespacio vectorial de $\mathbf{E}$.

DEMOSTRACIÓN ■ 1. $\vec{0}_{\mathbf{E}} = 0 \cdot \vec{u}_1 + 0 \cdot \vec{u}_2 + \cdots + 0 \cdot \vec{u}_k \in S$.

2. Si $\vec{u}, \vec{v} \in S$, entonces existen $a_1, a_2, \ldots, a_k, b_1, b_2, \ldots, b_k \in \mathbb{R}$ tales que:

$$\begin{aligned} \vec{u} &= a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_k\vec{u}_k, \\ \vec{v} &= b_1\vec{u}_1 + b_2\vec{u}_2 + \cdots + b_k\vec{u}_k. \end{aligned}$$

Luego,

$$\begin{aligned} \vec{u} + \vec{v} &= \sum_{i=1}^{k} a_i\vec{u}_i + \sum_{i=1}^{k} b_i\vec{u}_i \\ &= (a_1 + b_1)\vec{u}_1 + (a_2 + b_2)\vec{u}_2 + \cdots + (a_k + b_k)\vec{u}_k; \end{aligned}$$

de donde $\vec{u} + \vec{v} \in S$.

3. Si $\lambda \in \mathbb{R}$ y $\vec{u} \in S$, entonces, para ciertos a_i, $i = 1, 2, \ldots, k$,

$$\vec{u} = a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_k\vec{u}_k.$$

Luego,

$$\begin{aligned}\lambda\vec{u} &= \lambda(a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_k\vec{u}_k)\\ &= (\lambda a_1)\vec{u}_1 + (\lambda a_2)\vec{u}_2 + \cdots + (\lambda a_k)\vec{u}_k;\end{aligned}$$

por lo que $\lambda\vec{u} \in S$.

De 1, 2 y 3, $S < \mathbf{E}$. ■

Definición 3.12 ***(Subespacio generado por vectores)*** *Sean* **E** *un espacio vectorial y* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ *vectores en* **E**. *Al subespacio vectorial de* **E**,[10] *de todas las combinaciones lineales de estos vectores, se le denota como*

$$\operatorname{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k)$$

y se llama el **subespacio generado** *por los vectores* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$.

► **Ejemplo 3.27** Sea el espacio vectorial $S = \operatorname{gn}(\operatorname{sen}^2 x, \cos^2 x)$, determinar si $\cos(2x) \in S$. ◄

Solución $\cos(2x) = \cos^2 x - \operatorname{sen}^2 x = (1)\cos^2 x + (-1)\operatorname{sen}^2 x$. Por tanto, $\cos 2x \in S$. ✔

► **Ejemplo 3.28** ¿$2x^2 - 3x + 1 \in \operatorname{gn}(2, 1-x, x^2)$? ◄

Solución Encontremos $a_1, a_2, a_3 \in \mathbb{R}$ tales que

$$a_1 \cdot 2 + a_2(1-x) + a_3(x^2) = 2x^2 - 3x + 1\,.$$

Entonces, al reducir,

$$(2a_1 + a_2) - a_2 x + a_3 x^2 = 2x^2 - 3x + 1.$$

Y puesto que dos polinomios son iguales si los coeficientes de las mismas potencias de x son iguales,[11] se tiene

$$\begin{aligned}2a_1 + a_2 &= 1,\\ -a_2 &= -3,\\ a_3 &= 2;\end{aligned}$$

de donde $a_1 = -1, a_2 = 3, a_3 = 2$. Por tanto, $2x^2 - 3x + 1 \in \operatorname{gn}(2, 1-x, x^2)$. ✔

► **Ejemplo 3.29** Determinar qué subespacio de las matrices cuadradas de orden 2 es

$$\operatorname{gn}\left(\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right)$$ ◄

[10] Por el teorema 3.6 este conjunto es efectivamente un subespacio vectorial.

[11] Dos polinomios $p(x) = \sum_{k=0}^{m} a_k x^k$ y $q(x) = \sum_{k=0}^{m} b_k x^k$ son iguales si $a_k = b_k$ para cada $k = 0, 1, \ldots, m$.

Solución Toda matriz en este subespacio tiene la forma

$$a\begin{bmatrix}1&0\\0&0\end{bmatrix}+c\begin{bmatrix}0&1\\1&0\end{bmatrix}+d\begin{bmatrix}0&0\\0&1\end{bmatrix}=\begin{bmatrix}a&c\\c&d\end{bmatrix}, a,c,d\in\mathbb{R},$$

que es una matriz simétrica. Es fácil probar que, inversamente, toda matriz simétrica 2×2 es de esta forma (pruébelo). Luego, el espacio generado por estas matrices es el subespacio de las matrices simétricas cuadradas de orden 2 (cfr. ejemplo 3.25). ✔

► **Ejemplo 3.30** En **P**, el espacio de polinomios, si $S=\text{gn}(1,x,x^2,x^3)$, entonces $S=\mathbf{P}_3$ (polinomios de grado a lo más tres y el polinomio constante cero); pues todo elemento de $\mathbf{P}_3$ tiene la forma

$$a_0+a_1x+a_2x^2+a_3x^3=a_0\cdot 1+a_1\cdot x+a_2\cdot x^2+a_3\cdot x^3\in S,$$

para ciertos $a_i\in\mathbb{R}$, $i=0,1,2,3$.◄

Espacio fila y espacio columna de una matriz

Definición 3.13 *Si $A\in\mathfrak{M}_{m\times n}$, se define el* **espacio fila** *de A como el subespacio de $\mathbb{R}^n$ generado por las m-filas de A; y al subespacio de $\mathbb{R}^m$, generado por las n-columnas de A, se le llama el* **espacio columna** *de A. Al espacio fila y al espacio columna se les denotará, en este libro, como $E_f(A)$ y $E_c(A)$, respectivamente.*

► **Ejemplo 3.31** Si $\begin{bmatrix}3&-1&2&0\\2&-1&3&4\\0&1&2&3\end{bmatrix}\in\mathfrak{M}_{3\times 4}$,

$$\begin{aligned}E_f(A)&=\text{gn}((3,-1,2,0),(2,-1,3,4),(0,1,2,3)),\\E_c(A)&=\text{gn}((3,2,0),(-1,-1,1)(2,3,2),(0,4,3)).\end{aligned}$$◄

Teorema 3.7 *Sea $A\in\mathfrak{M}_{m\times n}$, entonces $\vec{b}\in E_c(A)\Leftrightarrow$ el sistema $A\vec{x}=\vec{b}$ tiene solución.*

DEMOSTRACIÓN ■ Supongamos que

$$A=\begin{bmatrix}a_{11}&a_2&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}.$$

Entonces $\vec{b}\in E_c(A)$ si y sólo si existen $x_1,x_2,\ldots,x_n\in\mathbb{R}$ tales que

$$\vec{b}=x_1\begin{bmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{bmatrix}+x_2\begin{bmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{bmatrix}+\cdots+x_n\begin{bmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{bmatrix}$$

y la última igualdad se da si y sólo si[12]

$$\vec{b}=\begin{bmatrix} a_{11} & a_2 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix};$$

es decir, si y sólo si el sistema $A\vec{x}=\vec{b}$ tiene solución. ■

Criterio para determinar si un vector en $\mathbb{R}^n$ es combinación lineal de otros vectores

Podemos utilizar el teorema 3.7 para determinar si un vector dado, $\vec{b}$, es combinación lineal de los vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ en $\mathbb{R}^n$ (lo cual equivale a que $\vec{b}$ pertenezca al subespacio generado por estos vectores) de la siguiente manera:

1. Se forma la matriz cuyas columnas son los vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$.
2. Se resuelve el sistema $A\vec{x}=\vec{b}$ (por el método de Gauss, de preferencia).
3. Si este sistema tiene solución $\vec{x}=(a_1, a_2, \ldots, a_k)$, entonces $\vec{b}$ pertenece al subespacio generado por los vectores $\vec{u}_i$ (es combinación lineal de ellos), y además

$$\vec{b}=a_1\vec{u}_1+a_2\vec{a}_2+\cdots+a_k\vec{u}_k.$$

4. Si el sistema $A\vec{x}=\vec{b}$ es inconsistente, entonces $\vec{b}$ no es combinación lineal de los vectores $\vec{u}_i$ ($\vec{b}\notin \text{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k)$).

► **Ejemplo 3.32** Determinar si $\vec{b}=(2,6,1)$ pertenece al espacio generado por $\vec{u}_1=(1,2,3)$, $\vec{u}_2=(-2,-5,2)$, $\vec{u}_3=(1,2,-1)$. En caso afirmativo, hallar a_1, a_2, a_3, tales que $\vec{b}=a_1\vec{u}_1+a_2\vec{u}_2+a_3\vec{u}_3$.◄

Solución Resolvamos el sistema

$$\begin{bmatrix} 1 & -2 & 1 \\ 2 & -5 & 2 \\ 3 & 2 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 2 \\ 6 \\ 1 \end{bmatrix}$$

por el método de Gauss:

$$\left[\begin{array}{rrr|r} 1 & -2 & 1 & 2 \\ 2 & -5 & 2 & 6 \\ 3 & 2 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & -2 & 1 & 2 \\ 0 & -1 & 0 & 2 \\ 0 & 8 & -4 & -5 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & -2 & 1 & 2 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & -4 & 11 \end{array}\right].$$

Como el sistema es consistente, $\vec{b}\in \text{gn}(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ y, al hacer sustitución regresiva, $a_3=-11/4$, $a_2=-2$, $a_1=3/4$; esto es,

$$\vec{b}=(3/4)\vec{u}_1-2\vec{u}_2+(-11/4)\vec{u}_3. \quad ✔$$

[12] Cfr. (1.5) del ejemplo 1.15, de la página 12.

Generadores de un espacio

Definición 3.14 *Sean* **E** *un espacio vectorial y* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k \in \mathbf{E}$. *Los vectores* $\vec{u}_1$, $\vec{u}_2$, $\ldots$, $\vec{u}_k$ **generan** *a* **E** *si todo elemento de* **E** *es combinación lineal de estos vectores; i.e.,* $\mathbf{E} = \operatorname{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k)$. *En tal caso, se dice que los vectores* $\vec{u}_i$ *forman un conjunto de generadores para este espacio vectorial.*

▶ **Ejemplo 3.33** Sea $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathfrak{M}_2$ cualquier matriz. Entonces

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Luego

$$\mathfrak{M}_2 = \operatorname{gn}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right).$$ ◀

▶ **Ejemplo 3.34** (Generadores del espacio $\mathfrak{M}_{m\times n}$). En general, si $\mathfrak{M}_{m\times n}$ es el espacio vectorial de las matrices de tamaño $m \times n$, entonces las mn matrices definidas en este espacio como $M_{\alpha\beta} = [m_{ij}]$, $1 \leq \alpha \leq m$, $1 \leq \beta \leq n$, donde

$$m_{ij} = \begin{cases} 1 & \text{si } i = \alpha,\ j = \beta, \\ 0 & \text{en otro caso;} \end{cases}$$

(por ejemplo $M_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ en las matrices de tamaño 2×3) generan al espacio[13] $\mathfrak{M}_{m\times n}$. Lo cual es fácil de probar y se deja como ejercicio al lector. ◀

▶ **Ejemplo 3.35** Sean los vectores $\vec{e}_i$ del espacio $\mathbb{R}^n$ definidos como

$$\vec{e}_i = (\underbrace{0, \ldots, 0, \overset{i}{1}, 0, \ldots, 0}_{n})$$

para $i = 1, 2, \ldots, n$. Entonces, si $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, se tiene

$$\vec{u} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n.$$

Así que los vectores $\vec{e}_i$, $i = 1, 2, \ldots, n$, generan a $\mathbb{R}^n$; esto es,

$$\mathbb{R}^n = \operatorname{gn}(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n).$$

Por ejemplo, si $n = 3$, $\mathbb{R}^3 = \operatorname{gn}((1,0,0),(0,1,0),(0,0,1))$. ◀

[13] De aquí en adelante supondremos que los espacios con los que se trabaje en este libro estarán dotados de las operaciones usuales que hemos definido en cada caso, a menos que se especifique lo contrario.

Criterios para que k vectores en $\mathbb{R}^n$ generen a $\mathbb{R}^n$

Supongamos que se tienen n vectores $\vec{u}_i$ de $\mathbb{R}^n$. Sabemos por el teorema 3.7, o del criterio que después de él se da (cfr. pág. 147), que todo vector $\vec{b} \in \mathbb{R}^n$ es combinación lineal de los vectores $\vec{u}_i$ si y sólo si el sistema $A\vec{x} = \vec{b}$, donde A es la matriz cuyas columnas son los vectores $\vec{u}_i$, tiene solución para todo $\vec{b} \in \mathbb{R}^n$. Pero, por el teorema 2.3 de la sección 2.1.2 (cfr. pág. 67), el sistema $A\vec{x} = \vec{b}$ tiene solución para todo $\vec{b}$ si y sólo si la matriz A es invertible.[14] Con esto hemos probado el siguiente teorema.

Teorema 3.8 *n-vectores de $\mathbb{R}^n$ generan a $\mathbb{R}^n$ si y sólo si la matriz que tiene a estos vectores como columnas es invertible.*

► **Ejemplo 3.36** Determinar si los vectores $(1,-2,3)$, $(2,-1,4)$ y $(-1,1,0)$ generan a $\mathbb{R}^3$.◄

Solución Por el teorema anterior, tres vectores generan a $\mathbb{R}^3$ si y sólo si la matriz que tiene a éstos como columnas es invertible; pero sabemos que una matriz es invertible si y sólo si es equivalente a la identidad (cfr. teorema 2.3, pág. 67). Llevemos esta matriz a la forma escalonada, por el método de Gauss, para ver si es o no equivalente a la identidad:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ -2 & -1 & 1 \\ 3 & 4 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 0 & -2 & 3 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 0 & 0 & 7 \end{bmatrix}.$$

Dado que la última matriz tiene pivote en todas las columnas, se sigue que A es equivalente a la identidad. Luego los vectores $(1,-2,3)$, $(2,-1,4)$ y $(-1,1,0)$ generan a $\mathbb{R}^3$. ✔

Supongamos ahora que tenemos k vectores en $\mathbb{R}^n$ con $k < n$ (por ejemplo tres vectores en $\mathbb{R}^4$). Para que estos vectores generen a $\mathbb{R}^n$, por el criterio de la página 148, es necesario y suficiente que el sistema $A\vec{x} = \vec{b}$ tenga solución para todo $\vec{b} \in \mathbb{R}^n$, donde A es la matriz que tiene como columnas a estos vectores. Sin embargo, al llevar A a su forma escalonada reducida, puesto que $k < n$, la última fila de esta forma escalonada reducida tiene necesariamente todas sus componentes nulas (piense por ejemplo, para fijar ideas, en una matriz 3×2). Sea F_i la fila de la matriz A que se transformó en la última fila nula de la forma escalonada reducida de la matriz A; tomemos el vector $\vec{b} = (0,\ldots,0,\overset{i}{1},0,\ldots,0) \in \mathbb{R}^n$, entonces el sistema $A\vec{x} = \vec{b}$ será inconsistente. Luego $\vec{b}$ no puede ser combinación lineal de los vectores columna de A. Así que dichos vectores no generan a $\mathbb{R}^n$. Con esta discusión hemos probado el siguiente teorema.

Teorema 3.9 *k vectores en $\mathbb{R}^n$, con $k < n$, no generan a $\mathbb{R}^n$.*

► **Ejemplo 3.37** Los vectores $(1,2,-1,4), (-1,2,1,0), (0,1,3,1)$ no generan a $\mathbb{R}^4$.◄

---

[14] Recuerde que, a su vez, una matriz es invertible si y sólo si es equivalente a la identidad.

Tenemos un tercero y último caso para que k vectores generen a $\mathbb{R}^n$ cuando $k > n$. Nuevamente, estos vectores generan este espacio si y sólo si el sistema $A\vec{x} = \vec{b}$ tiene solución para todo $\vec{b} \in \mathbb{R}^n$, donde A, como antes, es la matriz que tiene por columnas a dichos vectores. Pensemos en la forma escalonada reducida, H, de A y supongamos que tiene n pivotes (por tanto, cualquier forma escalonada equivalente a A tendrá también n pivotes). Entonces, puesto que $k > n$, esta forma escalonada no puede tener filas nulas; de aquí que el sistema $A\vec{x} = \vec{b}$ es consistente para todo $\vec{b}$. Además, dicho sistema tiene por tanto $k - n$ variables libres (las columnas donde no hay pivote) independientemente del vector $\vec{b}$, y dado que pueden tomar cualquier valor, en particular podemos elegir que todas sean cero. Por ende, todo vector $\vec{b}$ es combinación lineal de los vectores columna de la matriz A que corresponden a columnas con pivote[15] en H. Es decir, $\mathbb{R}^n$ está generado por los vectores columna de A que correspondan a columnas con pivote en cualquier forma escalonada equivalente a la matriz A. Por otra parte, si una forma escalonada (y, por tanto, cualquier otra forma escalonada $U \sim A$) equivalente a la matriz A no tiene n pivotes, entonces, dado que $k > n$ (piense por ejemplo en una matriz 3×4), esta forma escalonada equivalente tiene por lo menos una fila nula; y por el mismo argumento dado en las demostraciones de los teoremas 3.8 y 3.9, el sistema $A\vec{x} = \vec{b}$ es inconsistente para ciertos vectores $\vec{b}$, por lo que los vectores columna de A no generan a $\mathbb{R}^n$. Hemos probado así el siguiente teorema.

Teorema 3.10 *k-vectores de $\mathbb{R}^n$, con $k > n$, generan a $\mathbb{R}^n$ si y sólo si una forma escalonada equivalente de la matriz A, cuyas columnas son estos vectores, tiene n pivotes.*[16] *Si éste es el caso, podemos suprimir aquellos vectores de A que correspondan a columnas de la forma escalonada que no tengan pivote; los vectores restantes generarán a $\mathbb{R}^n$.*

► **Ejemplo 3.38** Sean $\vec{v}_1 = (1,2,0)$, $\vec{v}_2 = (1,1,1)$, $\vec{v}_3 = (1,4,-2)$, $\vec{v}_4 = (2,3,2)$.

1. Determinar si los vectores $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ generan a $\mathbb{R}^3$.
2. En caso afirmativo, reducir a un mínimo el número de generadores.◄

Solución 1. Formamos la matriz que tiene como columnas a estos vectores y la llevamos a la forma escalonada:

$$\begin{bmatrix} 1 & 1 & 1 & 2 \\ 2 & 1 & 4 & 3 \\ 0 & 1 & -2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & -1 & 2 & -1 \\ 0 & 1 & -2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Como hay tres pivotes, $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ generan a $\mathbb{R}^3$.

2. De la matriz original, excluimos el vector columna que corresponde en la forma escalonada a la columna que no tiene pivote. Por tanto, $\mathbb{R}^3 = \operatorname{gn}(\vec{v}_1, \vec{v}_2, \vec{v}_4)$. ✔

[15]Por ejemplo, si las últimas $k-n$ variables son libres y $\vec{u}_i$ son los n primeros vectores columna de A, en este caso se tendría $\vec{b} = \sum_{i=1}^{n} x_i\vec{u}_i = A\vec{x}$, donde $\vec{x} = (x_1, x_2, \ldots, x_n, \overbrace{0, \ldots, 0}^{k-n})$ es la solución de este sistema para las variables libres que hemos elegido.

[16]Recuerde que todas las matrices en forma escalonada equivalentes a la matriz A tienen el mismo número de pivotes y que éstos se encuentran en las mismas posiciones (cfr. nota 1.5, pág. 29).

3.3 Dependencia e independencia lineal

Un concepto clave, estrechamente relacionado con el de combinación lineal, es el de dependencia e independencia lineal de vectores. En la figura 3-23 tenemos tres casos en el plano cartesiano. En *(a)*, un vector es múltiplo escalar del otro y, por tanto, combinación lineal del primero. En *(b)* ningún vector se puede escribir como combinación lineal del otro. Y en *(c)* $\vec{w} = \alpha\vec{u} + \beta\vec{v}$; es decir, uno de los vectores es combinación lineal de los otros. Se dice entonces que los vectores en *(a)* y *(c)* son linealmente dependientes, mientras que en *(b)* son linealmente independientes.

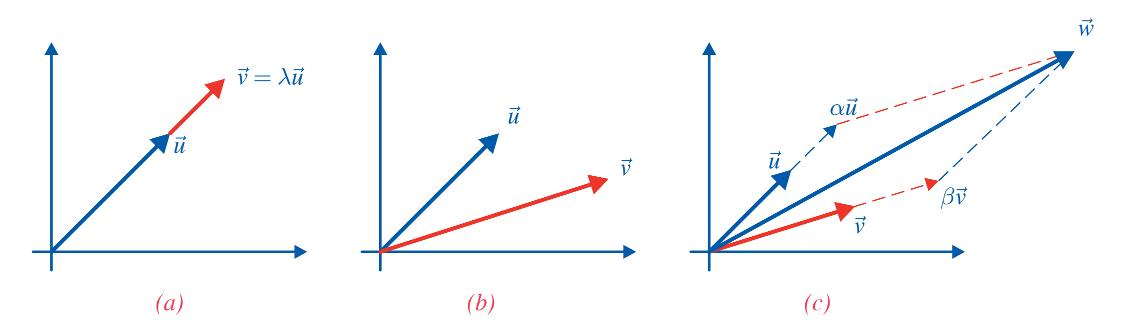


Figura 3-23 • Tres casos representativos de dependencia lineal en $\mathbb{R}^2$: *(a)* vectores linealmente dependientes; *(b)* vectores linealmente independientes, y *(c)* vectores linealmente dependientes.

Generalizamos el concepto de dependencia e independencia lineal a cualquier espacio vectorial en la siguiente definición:

Definición 3.15 *Sean* $\mathbf{E}$ *un espacio vectorial y* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$, *k vectores en* $\mathbf{E}$. *Se dice que estos vectores son* **linealmente dependientes** *(L.D.) si uno de ellos es combinación lineal de los otros. En caso contrario diremos que los vectores son* **linealmente independientes** *(L.I.).*[17]

Supongamos que los vectores $\vec{u}_i \in \mathbf{E}$, $i = 1, 2, \ldots, k$, son linealmente dependientes. Entonces uno de ellos es combinación lineal de los otros, digamos $\vec{u}_i$. Por tanto, existen escalares α_j tales que

$$\vec{u}_i = \alpha_1\vec{u}_1 + \cdots + \alpha_{i-1}\vec{u}_{i-1} + \alpha_{i+1}\vec{u}_{i+1} + \cdots + \alpha_k\vec{u}_k;$$

de aquí que

$$\alpha_1\vec{u}_1 + \cdots + \alpha_{i-1}\vec{u}_{i-1} + (-1)\,\vec{u}_i + \alpha_{i+1}\vec{u}_{i+1} + \cdots + \alpha_k\vec{u}_k = \vec{0}_{\mathbf{E}}$$

y uno, por lo menos, de los coeficientes de los $\vec{u}_j$ es distinto de cero (específicamente el coeficiente de $\vec{u}_i$ que es -1). Inversamente, supongamos que existen escalares α_j, $j = 1, 2, \ldots k$, con uno de ellos distinto de cero, tales que

$$\alpha_1\vec{u}_1 + \cdots + \alpha_{i-1}\vec{u}_{i-1} + \alpha_i\vec{u}_i + \alpha_{i+1}\vec{u}_{i+1} + \cdots + \alpha_k\vec{u}_k = \vec{0}_{\mathbf{E}}.$$

[17] Abreviaremos linealmente independientes como L.I. y linealmente dependientes como L.D. a lo largo de todo este libro.

Digamos que α_i es uno de estos coeficientes no nulos; entonces

$$\vec{u}_i = -(\alpha_1/\alpha_i)\vec{u}_1 - \cdots - (\alpha_{i-1}/\alpha_i)\vec{u}_{i-1} - (\alpha_{i+1}/\alpha_i)\vec{u}_{i+1} - \cdots - (\alpha_k/\alpha_i)\vec{u}_k.$$

Luego $\vec{u}_i$ es combinación lineal de los restantes vectores; por tanto, son vectores linealmente dependientes. Con esta discusión hemos hecho patente la demostración del siguiente teorema que establece condiciones equivalentes muy útiles para dependencia e independencia lineal.

Teorema 3.11 *Sean $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ vectores en un espacio vectorial* **E**. *Entonces:*

1. *Los vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ son* ***linealmente dependientes*** *si y sólo si existen $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$, con uno de ellos (por lo menos) distinto de cero, tales que*

$$\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_k\vec{u}_k = \vec{0}_{\mathbf{E}}.$$

2. *Los vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ son* ***linealmente independientes*** *si y sólo si los únicos escalares $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$ tales que*

$$\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_k\vec{u}_k = \vec{0}_{\mathbf{E}}$$

son $\alpha_1 = 0, \alpha_2 = 0, \ldots, \alpha_k = 0$.

❍ **Nota 3.10**

1. Observe que si los vectores $\vec{u}_i$, $i = 1, 2, \ldots, k$, son L.I., entonces cualquier subconjunto no vacío de ellos es también L.I.
2. Los vectores $\vec{u}_i$, $i = 1, 2, \ldots, k$, son L.D. si y sólo si hay un subconjunto de ellos que es L.D.
3. Si alguno de los vectores $\vec{u}_i$, $i = 1, 2, \ldots, k$, es el vector neutro aditivo del espacio, $\vec{0}_{\mathbf{E}}$, entonces estos vectores son L.D.; en particular, cuando $k = 1$, $\vec{u}$ es L.I. si y sólo si $\vec{u} \neq \vec{0}_{\mathbf{E}}$.

► **Ejemplo 3.39** ¿Son los vectores $\vec{v}_1 = (1,2,3,1)$, $\vec{v}_2 = (2,2,1,3)$ y $\vec{v}_3 = (-1,2,7,-3)$ linealmente dependientes o linealmente independientes?◄

Solución Si $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ y $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \alpha_3\vec{v}_3 = \vec{0}_{\mathbf{R^4}}$, entonces

$$(\alpha_1 + 2\alpha_2 - \alpha_3, 2\alpha_1 + 2\alpha_2 + 2\alpha_3, 3\alpha_1 + \alpha_2 + 7\alpha_3, \alpha_1 + 3\alpha_2 - 3\alpha_3) = (0,0,0,0),$$

que equivale al sistema

$$\begin{aligned} \alpha_1 + 2\alpha_2 - \alpha_3 &= 0 \\ 2\alpha_1 + 2\alpha_2 + 2\alpha_3 &= 0 \\ 3\alpha_1 + \alpha_2 + 7\alpha_3 &= 0 \\ \alpha_1 + 3\alpha_2 - 3\alpha_3 &= 0. \end{aligned}$$

Resolvamos ahora este sistema homogéneo:

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 2 \\ 3 & 1 & 7 \\ 1 & 3 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & -2 & 4 \\ 0 & -5 & 10 \\ 0 & -1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 0 & -1 & 2 \\ 0 & -1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Por tanto $\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} -3r \\ 2r \\ r \end{bmatrix}$ $r \in \mathbb{R}$; es decir, el sistema tiene solución no trivial; así, por ejemplo, con $r = 1$, $-3\vec{v}_1 + 2\vec{v}_2 + \vec{v}_3 = \vec{0}$. Luego $\vec{v}_1, \vec{v}_2, \vec{v}_3$ son linealmente dependientes. ✔

► **Ejemplo 3.40** Determinar si $x^2 - 1, x^2 + 1, 4x, 2x - 3$ son linealmente dependientes o linealmente independientes en el espacio de los polinomios.◄

Solución Si

$$\alpha_1(x^2 - 1) + \alpha_2(x^2 + 1) + \alpha_3(4x) + \alpha_4(2x - 3) = 0,$$

entonces

$$(\alpha_1 + \alpha_2)x^2 + (4\alpha_3 + 2\alpha_4)x - \alpha_1 + \alpha_2 - 3\alpha_4 = 0.$$

De donde obtenemos el sistema homogéneo

$$\begin{aligned} \alpha_1 + \alpha_2 &= 0 \\ 4\alpha_3 + 2\alpha_4 &= 0 \\ -\alpha_1 + \alpha_2 - 3\alpha_4 &= 0. \end{aligned}$$

Llevemos a forma escalonada la matriz de coeficientes:

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 4 & 2 \\ -1 & 1 & 0 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 2 & 0 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & -3 \\ 0 & 0 & 2 & 1 \end{bmatrix}.$$

El sistema tiene infinidad de soluciones dadas por

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = r \begin{bmatrix} -3/2 \\ 3/2 \\ -1/2 \\ 0 \end{bmatrix} = \begin{bmatrix} -3s \\ 3s \\ -s \\ 2s \end{bmatrix}.$$

En particular, tiene soluciones no triviales, y por ende los polinomios son linealmente dependientes. ✔

► **Ejemplo 3.41** ¿Son las matrices $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ L.I. o L.D. en el espacio de las matrices simétricas?◄

Solución Resolvamos el sistema

$$\alpha_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

que equivale a

$$\begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_2 & \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix};$$

de donde $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Luego estas matrices son L.I. ✔

▶ **Ejemplo 3.42** Determinar si las funciones $f_1(x) = 1$, $f_2(x) = x$ y $f_3(x) = x^2$ son L.I. o L.D. en el espacio de funciones $\mathscr{F}(-\infty,\infty)$ (las funciones con dominio en el intervalo $(-\infty,\infty)$).◀

Solución Sean $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ tales que

$$\alpha_1 f_1 + \alpha_2 f_2 + \alpha_3 f_3 = \theta.$$

Entonces tenemos

$$\alpha_1 + \alpha_2 x + \alpha_3 x^2 = 0 \quad \forall x \in \mathbb{R}.$$

En particular, para $x = 0$, $x = 1$ y $x = -1$; por lo que

$$\begin{aligned} \alpha_1 &= 0, \\ \alpha_2 + \alpha_3 &= 0, \\ -\alpha_2 + \alpha_3 &= 0; \end{aligned}$$

de donde $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Por tanto, f_1, f_2, f_3 son L.I. en este espacio. ✔

▶ **Ejemplo 3.43** Determinar si las funciones f_1 y f_2 son linealmente dependientes o independientes en el espacio $C[-1,1]$ (el espacio de funciones continuas en el intervalo $[-1,1]$; cfr. el ejemplo 3.23), donde

$$f_1(x) = \begin{cases} 0 & \text{si} \quad -1 \le x < 0 \\ x^2 & \text{si} \quad 0 \le x \le 1 \end{cases}$$

y

$$f_2(x) = \begin{cases} x^2 & \text{si} \quad -1 \le x < 0 \\ 0 & \text{si} \quad 0 \le x \le 1 \end{cases}$$

Las gráficas de estas funciones se encuentran en la figura 3-24.◀

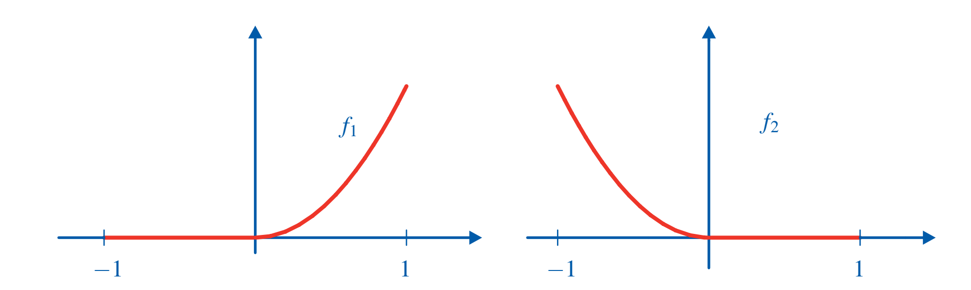


Figura 3-24 • Un par de funciones continuas linealmente independientes.

Solución Nuevamente, si $\alpha_1 f_1 + \alpha_2 f_2 = \theta$, donde θ es la función constante cero (el neutro aditivo de este espacio), entonces

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) = \theta(x) = 0 \quad \forall x \in [-1,1].$$

En particular si $x=-1$, se tiene

$$\alpha_1\cdot 0+\alpha_2\cdot 1=0;$$

y si $x=1$,

$$\alpha_1\cdot 1+\alpha_2\cdot 0=0.$$

De donde $\alpha_1=\alpha_2=0$; por tanto f_1 y f_2 son L.I. ✔

► **Ejemplo 3.44** (El wronskiano). Sea J un intervalo abierto y supongamos que $f_1, f_2, \ldots, f_n$ son funciones L.D. en el espacio $\mathscr{F}(J)$ y que todas estas funciones tienen derivada hasta el orden $n-1$ en todo punto de J, entonces existen constantes $\alpha_1, \alpha_2, \ldots, \alpha_n\in\mathbb{R}$ con una de ellas, al menos, distinta de cero, tales que

$$\alpha_1 f_1(x)+\alpha_2 f_2(x)+\cdots+\alpha_n f_n(x)=0\ \forall x\in J.$$

Así, para todo $x\in J$,

$$\begin{array}{ccccccccc}
\alpha_1 f_1(x) & + & \alpha_2 f_1(x) & + & \cdots & + & \alpha_n f_n(x) & = & 0\\
\alpha_1 f_1'(x) & + & \alpha_2 f_1'(x) & + & \cdots & + & \alpha_n f_n'(x) & = & 0\\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots\\
\alpha_1 f_1^{(n-1)}(x) & + & \alpha_2 f_2^{(n-1)}(x) & + & \cdots & + & \alpha_n f_n^{(n-1)}(x) & = & 0.
\end{array}\tag{3.12}$$

Es decir,

$$\begin{bmatrix} f_1(x) & f_2(x) & \cdots & f_n(x)\\ f_1'(x) & f_2'(x) & \cdots & f_2'(x)\\ \vdots & \vdots & \ddots & \vdots\\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x)\end{bmatrix}\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}=\begin{bmatrix}0\\0\\ \vdots\\0\end{bmatrix}.$$

Como una de las $\alpha_i\neq 0$, el sistema homogéneo (3.12) tiene solución no trivial y entonces, para todo $x\in J$,

$$\begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x)\\ f_1'(x) & f_2'(x) & \cdots & f_2'(x)\\ \vdots & \vdots & \ddots & \vdots\\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x)\end{vmatrix}=0.$$

Al determinante del lado izquierdo de la precedente igualdad se le llama el wronskiano de las funciones f_1, f_2, $\ldots$, f_n y se le denota por $W(f_1, f_2, \ldots, f_n)(x)$. Así, hemos probado que si las funciones son linealmente dependientes en J, es decir, $\mathscr{F}(J)$ para ser precisos, el wronskiano es cero en todo punto del intervalo J. Luego, si existe $x_0\in J$ tal que $W(f_1, f_2, \ldots, f_n)(x_0)\neq 0$, entonces las funciones son L.I. Sin embargo, si $W(f_1, f_2, \ldots, f_n)(x)=0$ para todo $x\in J$, las funciones no son necesariamente L.I. en J como lo hace patente el ejemplo 3.43.◄

► **Ejemplo 3.45** Los polinomios $1, x, \ldots, x^k$, son L.I. en el espacio de polinomios. En efecto, si $\alpha_i\in\mathbb{R}$ son tales que $\sum_{i=1}^k \alpha_i x^i=0$, entonces, por igualdad[18] de polinomios, se tiene $\alpha_i=0\ \forall i=1,2,\ldots,k$.◄

[18] Cfr. la nota al pie 11 de la página 145 y el ejemplo 3.28. En este caso, el polinomio constante 0 se puede escribir como $\sum_{i=1}^k 0\cdot x^k$.

3.3.1 Criterios de independencia lineal en $\mathbb{R}^n$

Sean $\vec{u}_i$, $i = 1,2,\ldots,k$, vectores en $\mathbb{R}^n$; estos vectores son L.D. si y sólo si el sistema

$$x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_k\vec{u}_k = \vec{0}_{\mathbb{R}^n}$$

tiene soluciones no triviales. Esto es, si el sistema homogéneo

$$A\vec{x} = \vec{0},$$

donde A es la matriz que tiene como columnas a los vectores $\vec{u}_i$, tiene soluciones no triviales. Del teorema 1.6 (cfr. pág. 32) sabemos que si $k = n$, este sistema tiene soluciones no triviales si y sólo si la matriz es no invertible. Y del mismo teorema sabemos que si $k > n$ (tenemos entonces un sistema con más variables que ecuaciones), el sistema homogéneo tiene soluciones no triviales. Con esto hemos probado el siguiente teorema.

Teorema 3.12 ***(Criterios para que k vectores de $\mathbb{R}^n$ sean L.I. o L.D.)*** *Sean $\vec{u}_1$, $\vec{u}_2,\ldots,\vec{u}_k$, k vectores en $\mathbb{R}^n$ y A la matriz que tiene como columnas a estos vectores. Entonces:*

1. *Los vectores son L.I. si y sólo si el sistema $A\vec{x} = \vec{0}$ tiene únicamente la solución trivial.*
2. *Los vectores son L.D. si y sólo si el sistema $A\vec{x} = \vec{0}$ tiene soluciones no triviales.*
3. *Si $k = n$:*
   - *(a) los vectores son L.I. si y sólo si A es invertible.*
   - *(b) los vectores son L.D. si y sólo si A no es invertible.*
4. *Si $k > n$ los vectores son L.D.*

► **Ejemplo 3.46** Determinar si los vectores dados son L.I. o L.D. en $\mathbb{R}^4$:

1. $(1,-2,-3,2)$, $(-1,0,1,1)$, $(3,2,1,1)$.
2. $(1,-2,-3,2)$, $(-1,0,1,1)$, $(5,-4,-9,1)$.
3. $(1,2,1,3)$, $(1,1,1,1)$, $(-1,1,0,0)$, $(-1,1,0,0)$, $(1,0,0,4)$.
4. $(1,0,1,0)$, $(2,-1,1,1)$, $(-2,1,1,3)$, $(-3,0,0,1)$.◄

Solución 1. Resolvamos el sistema homogéneo $A\vec{x} = \vec{0}$, donde A es la matriz que tiene como columnas a estos vectores.

$$\begin{bmatrix} 1 & -1 & 3 \\ -2 & 0 & 2 \\ -3 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 3 \\ 0 & -2 & 8 \\ 0 & -2 & 10 \\ 0 & 3 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 3 \\ 0 & 1 & -4 \\ 0 & -2 & 10 \\ 0 & 3 & -5 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -1 & 3 \\ 0 & 1 & -4 \\ 0 & 0 & 2 \\ 0 & 0 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 3 \\ 0 & 1 & -4 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

De donde el sistema sólo tiene la solución trivial, por lo que los vectores son L.I.

2. Resolvamos el sistema homogéneo $A\vec{x} = \vec{0}$, donde A es la matriz que tiene como columnas a estos vectores.

$$\begin{bmatrix} 1 & -1 & 5 \\ -2 & 0 & -4 \\ -3 & 1 & -9 \\ 2 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 5 \\ 0 & -2 & 6 \\ 0 & -2 & 6 \\ 0 & 3 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 5 \\ 0 & 1 & -3 \\ 0 & -2 & -6 \\ 0 & 3 & -9 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -1 & 5 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

De donde el sistema tiene una infinidad de soluciones, a saber, haciendo sustitución regresiva,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2r \\ 3r \\ r \end{bmatrix}.$$

Por tanto tiene soluciones no triviales, así que los vectores son L.D.; una solución es, por ejemplo, $(x_1, x_2, x_3) = (-2, 3, 1)$, que se obtiene al hacer $r = 1$.

(Comprobación: $\quad -2(1,-2,-3,2) + 3(-1,0,1,1) + (5,-4,-9,1) = (0,0,0,0)$.)

3. En este caso tenemos 5 vectores en $\mathbb{R}^4$, por el teorema 3.12 son L.D. ($k > n$).

4. En este caso tenemos 4 vectores en $\mathbb{R}^4$, y serán L.I. si y sólo si la matriz A, que tiene de columnas a estos vectores, es invertible. Llevemos esta matriz a la forma escalonada:

$$A = \begin{bmatrix} 1 & 2 & -2 & -3 \\ 0 & -1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -2 & -3 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 3 & 3 \\ 0 & 1 & 3 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -2 & -3 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -2 & -3 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & -5 \end{bmatrix}.$$

Dado que toda columna tiene pivote en la última matriz equivalente, se sigue que A es equivalente a la identidad y, por tanto, es invertible; luego los vectores son L.I. ✔

Como mencionamos anteriormente, y el lector seguramente ya lo notó, los conceptos de independencia lineal y combinaciones lineales están estrechamente relacionados. En $\mathbb{R}^n$ esta relación se hace aún más patente por la herramienta que proveen las técnicas matriciales y la teoría de sistemas lineales que estudiamos en el capítulo 1. Por el teorema 3.8 (cfr. pág. 149), n vectores de $\mathbb{R}^n$ generan a $\mathbb{R}^n$ si y sólo si la matriz que tiene por columnas a estos vectores es invertible; a su vez, por el teorema 3.12 (cfr. pág. 156), esta matriz es invertible si y sólo si los n vectores son L.I. Hemos probado así el siguiente teorema que relaciona los conceptos de independencia lineal y el número de generadores para el espacio $\mathbb{R}^n$.

Teorema 3.13 *Sean $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$, n vectores de $\mathbb{R}^n$ y A la matriz que tiene por columnas a estos vectores, entonces las siguientes condiciones son equivalentes:*

1. *$\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ son L.I.*
2. *$\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ generan a $\mathbb{R}^n$.*
3. *A es invertible.*[19]

► **Ejemplo 3.47**

1. En el ejemplo 3.46 vimos que la matriz que tiene por columnas los vectores $(1,0,1,0)$, $(2,-1,1,1)$, $(-2,1,1,3)$, $(-3,0,0,1)$ de $\mathbb{R}^4$ es equivalente a la identidad y, por ende, es invertible; así que estos vectores son L.I. y generan a $\mathbb{R}^4$.
2. Sean $(1,0,1)$, $(-2,1,1)$ y $(-4,3,5)$; llevemos la matriz A, que tiene por columnas a estos vectores, a la forma escalonada:

$$\begin{bmatrix} 1 & -2 & -4 \\ 0 & 1 & 3 \\ 1 & 1 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -4 \\ 0 & 1 & 3 \\ 0 & 3 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -4 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}.$$

Puesto que la última columna en la última matriz equivalente no tiene pivote, la matriz A no es equivalente a la identidad y, por tanto, no es invertible; así que estos vectores no generan a $\mathbb{R}^3$ y son L.D.◄

3.4 Bases y dimensión

En la sección anterior vimos cómo espacios vectoriales pueden ser generados por ciertos conjuntos de vectores; es decir, que todo elemento de estos espacios se puede escribir como una combinación lineal de los vectores de esos conjuntos generadores. También ilustramos, con algunos casos, cómo es posible reducir a un mínimo los generadores (cfr. ejemplo 3.38 y teorema 3.10). El tema de minimizar el conjunto de generadores está relacionado con la independencia lineal de los vectores y con el concepto, que estudiaremos en la presente sección, de bases de espacios vectoriales.

3.4.1 Definiciones y ejemplos

Comenzamos extendiendo el concepto de dependencia e independencia lineal a cualquier subconjunto de vectores, aun si tiene una infinidad de elementos.

Definición 3.16 *Sea* **E** *un espacio vectorial y $S \subset$* **E**.

1. *S es* **linealmente independiente** *(L.I.), si los vectores de todo subconjunto finito de S son linealmente independientes en el sentido de la definición*[20] *3.15.*
2. *S es* **linealmente dependiente** *(L.D.) si contiene (al menos) un subconjunto finito de vectores linealmente dependientes en el sentido de la definición 3.15.*

[19]Recuerde que una matriz es invertible si y sólo si es equivalente por filas a la identidad; y que para mostrar que una matriz es equivalente a la identidad basta llevarla, por el método de Gauss, a la forma escalonada y verificar que toda columna (y por ende toda fila) tenga pivote.

[20]Cfr. página 151.

▶ **Ejemplo 3.48** En **P**, el espacio de polinomios, sea

$$S_1 = \{1, x, x^2, \dots, x^n, \dots\}.$$

S_1 es L.I., mientras que

$$S_2 = \{1, x, 2x, 2x^2, 3x^3, 4x^4, 5x^5, \dots\}$$

es L.D. En efecto, sea A un subconjunto finito de S_1; entonces los elementos de A tienen la forma (dado que A es finito) x^{m_i}, para k enteros m_i, $i = 1, 2, \dots, k$, con $0 \le m_1 < m_2 < \cdots < m_k$; luego estos vectores forman parte de los vectores $1, x, \dots, x^{m_i}, \dots, x^{m_k}$ que son L.I. (cfr. el ejemplo 3.45); por tanto, los polinomios en A son también L.I. (cfr. la nota 3.10, pág. 152). Es decir, todo subconjunto finito de S_1 es L.I.; por tanto S_1 es L.I. En cuanto a S_2, los polinomios x y $2x$ son claramente L.D.; por tanto S_2 es L.D.◀

Definición 3.17 *Sean* **E** *un espacio vectorial y* $S \subset \mathbf{E}$. *Se dice que* S *genera a* **E** *si todo elemento de* **E** *es combinación lineal de elementos de* S*; es decir, para cada* $\vec{u} \in \mathbf{E}$ *existen* $\vec{s}_1, \dots, \vec{s}_k \in S$ *y escalares* $a_1, \dots, a_k$ *tales que* $\vec{u} = a_1\vec{s}_1 + \cdots + a_k\vec{s}_k$.

▶ **Ejemplo 3.49** $S = \{1, x, x^2, \dots, x^n, \dots\}$ genera a **P**. Pues claramente todo polinomio de este espacio tiene la forma $a_0 \cdot 1 + a_2 \cdot x + \cdots + a_k \cdot x^k$.◀

Definición 3.18 ***(Bases)*** *Si* **E** *es un espacio vectorial y* $\mathscr{B} \subset \mathbf{E}$, $\mathscr{B}$ *es una* **base** *de* **E** *si:*

1. $\mathscr{B}$ *es L.I.*
2. $\mathscr{B}$ *genera a* **E**.

▶ **Ejemplo 3.50** (Base canónica de $\mathbb{R}^n$). En $\mathbb{R}^n$ sean los vectores

$$\vec{e}_i = (\underbrace{0, \dots, 0, \overset{i}{1}, 0, \dots, 0}_{n})$$

$i = 1, 2, \dots, n$. Entonces

$$\mathscr{B} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$$

es una base (la llamada base *canónica* o *estándar*) de $\mathbb{R}^n$; pues la matriz que tiene por columnas a estos vectores es la identidad $n \times n$ que es invertible, y por el teorema 3.13 los vectores generan a $\mathbb{R}^n$ y son L.I.◀

▶ **Ejemplo 3.51** (Base de **P**). En **P**, $\mathscr{B} = \{1, x, x^2, x^3, \dots, x^n, \dots\}$ es una base. Pues en el ejemplo 3.48 vimos que este conjunto es L.I., y en el ejemplo 3.49 mostramos que este conjunto genera a **P**.◀

▶ **Ejemplo 3.52** (Base de $\mathbf{P}_k$). En $\mathbf{P}_k$, el espacio de polinomios de grado a lo más k,

$$\mathscr{B} = \{1, x, x^2, x^3, \dots, x^k\}$$

es una base. Pues en el ejemplo 3.45 probamos que este conjunto es L.I. y claramente todo polinomio de $\mathbf{P}_k$ es de la forma $\sum_{i=0}^{k} a_i x^i$.◀

▶ **Ejemplo 3.53** (Base de $\mathfrak{M}_{m\times n}$). Sean las matrices $M_{\alpha\beta} = [m_{ij}]$ de tamaño $m \times n$ definidas en el ejemplo 3.34; es decir,

$$m_{ij} = \begin{cases} 1 & \text{si } i = \alpha, \beta = j, \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces, si $A = [a_{ij}] \in \mathfrak{M}_{m\times n}$, es claro que[21]

$$A = \sum_{\substack{1\leq\alpha\leq m \\ 1\leq\beta\leq n}} a_{\alpha\beta} M_{\alpha\beta} \tag{3.13}$$

lo cual entraña que las mn matrices $M_{\alpha\beta}$ generan a $\mathfrak{M}_{m\times n}$. Por otra parte, si el lado derecho de (3.13) se iguala a la matriz cero $m \times n$, entonces (por definición de la igualdad de matrices) se obtiene que las constantes $a_{\alpha\beta} = 0$, $\forall\alpha,\beta$; es decir, las matrices $M_{\alpha\beta}$ también son L.I. De esta manera las mn matrices $M_{\alpha\beta}$ forman una base para el espacio $\mathfrak{M}_{m\times n}$. Por ejemplo, en $\mathfrak{M}_{3\times 2}$,

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

es una base.◀

Enunciamos el siguiente teorema, cuya demostración es delicada y la postergamos al apéndice C. Invitamos al lector a que la consulte, pues contiene una construcción muy interesante que se utiliza con frecuencia en matemáticas para probar diversos teoremas de existencia y que depende de un principio muy básico y fundamental que es el llamado lema de Zorn.

Teorema 3.14 ***(Existencia de bases)*** *Sea* **E** *un espacio vectorial. Entonces* **E** *posee una base; es decir, existe* $\mathscr{B} \subset \mathbf{E}$ *L.I. tal que todo vector del espacio* **E** *es combinación lineal de elementos de* $\mathscr{B}$.

3.4.2 Dimensión, extracción de bases y compleción de un conjunto L.I. a una base

Hemos visto varios ejemplos de espacios vectoriales que se pueden generar mediante combinaciones lineales de elementos de ciertos subconjuntos de esos espacios. La idea central es optimizar el número de vectores generadores; y la independencia lineal es obviamente el discriminador natural para excluir vectores de un conjunto de generadores y quedarse con un mínimo de éstos que sigan generando al espacio. Las bases de espacios vectoriales son los conjuntos que cumplen con las condiciones de poder generar al espacio y ser L.I. Surge entonces la conjetura natural de que si dos bases generan al mismo espacio y sus elemementos son L.I., entonces tienen un mínimo de generadores; luego ambas deben tener el mismo número de elementos. Esta conjetura es cierta, y la probaremos en el teorema 3.16; lo cual da pauta a definir lo que llamaremos dimensión de un espacio vectorial. El concepto de dimensión de un espacio no es otra cosa que la generalización y abstracción de la idea intuitiva que tenemos de la dimen-

[21] La notación $\sum_{\substack{1\leq\alpha\leq m \\ 1\leq\beta\leq n}} a_{\alpha\beta}M_{\alpha\beta}$ significa que se suman todas las posibles combinaciones de $a_{\alpha\beta}M_{\alpha\beta}$ cuando los subíndices α y β recorren los valores $1,\ldots,m$ y $1,\ldots,n$, respectivamente. Dicho de paso, el número de términos de esta suma es mn.

sión de conjuntos que nos son familiares, como una línea recta, un plano y el espacio donde habitamos; a los cuales de manera natural les hemos asignado dimensiones uno, dos y tres, respectivamente, en el lenguaje coloquial.

Teorema 3.15 *Sea* **E** *un espacio vectorial que está generado por los vectores L.I.* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$. *Si* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ *son vectores en este espacio con* $k > n$, *entonces estos vectores son L.D.*

DEMOSTRACIÓN ■ Resolvamos el sistema

$$\beta_1 \vec{v}_1 + \beta_2 \vec{v}_1 + \cdots + \beta_k \vec{v}_k = \vec{0}_{\mathbf{E}} \tag{3.14}$$

para $\beta_i \in \mathbb{R}$, $1 = 1, 2, \ldots, k$. Dado que $\mathbf{E} = \operatorname{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$, existen escalares α_{ij}, $i = 1, \ldots, k$, $j = 1, \ldots, n$, tales que

$$\begin{aligned} \vec{v}_1 &= \alpha_{11}\vec{u}_1 + \alpha_{12}\vec{u}_2 + \cdots + \alpha_{1n}\vec{u}_n \\ \vec{v}_2 &= \alpha_{21}\vec{u}_1 + \alpha_{22}\vec{u}_2 + \cdots + \alpha_{2n}\vec{u}_n \\ &\vdots \\ \vec{v}_k &= \alpha_{k1}\vec{u}_1 + \alpha_{k2}\vec{u}_2 + \cdots + \alpha_{kn}\vec{u}_n \end{aligned} \tag{3.15}$$

Sustituyendo cada $\vec{v}_i$ de (3.15) en (3.14) obtenemos

$$\begin{aligned} \vec{0}_{\mathbf{E}} &= \beta_1 (\alpha_{11}\vec{u}_1 + \alpha_{12}\vec{u}_2 + \cdots + \alpha_{1n}\vec{u}_n) + \beta_2 (\alpha_{21}\vec{u}_1 + \alpha_{22}\vec{u}_2 + \cdots + \alpha_{2n}\vec{u}_n) \\ &\quad + \cdots + \beta_k (\alpha_{k1}\vec{u}_1 + \alpha_{k2}\vec{u}_2 + \cdots + \alpha_{kn}\vec{u}_n), \end{aligned}$$

que agrupando términos produce

$$\begin{aligned} \vec{0}_{\mathbf{E}} &= (\beta_1\alpha_{11} + \beta_2\alpha_{21} + \cdots + \beta_k\alpha_{k1})\vec{u}_1 + (\beta_1\alpha_{12} + \beta_2\alpha_{22} + \cdots + \beta_k\alpha_{k2})\vec{u}_2 \\ &\quad + \cdots + (\beta_1\alpha_{1n} + \beta_2\alpha_{2n} + \cdots + \beta_k\alpha_{kn})\vec{u}_n. \end{aligned}$$

Puesto que los vectores $\vec{u}_i$ son L.I., se debe tener

$$\begin{array}{ccccccccc} \beta_1\alpha_{11} & + & \beta_2\alpha_{21} & + & & + & \beta_k\alpha_{k1} & = & 0 \\ \beta_1\alpha_{12} & + & \beta_2\alpha_{22} & + & & + & \beta_k\alpha_{k2} & = & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta_1\alpha_{1n} & + & \beta_2\alpha_{2n} & + & \cdots & + & \beta_k\alpha_{kn} & = & 0, \end{array}$$

que es un sistema lineal homogéneo con más incógnitas que ecuaciones (recuerde que $k > n$); por tanto tiene solución no trivial; es decir, existen β_i, $i = 1, 2, \ldots, k$, no todos cero, tales que (3.14) se cumple. Luego los vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ son L.D. ■

El siguiente corolario es consecuencia inmediata del teorema 3.15 y su demostración se deja como ejercicio al lector.

Corolario 3.1 *Si un espacio vectorial* **E** *está generado por los vectores L.I.* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$, *entonces todo conjunto de vectores linealmente independiente tiene a lo más n vectores.*

Lema 3.2 *Sean* $\mathbf{E}$ *un espacio vectorial y* $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\}$ *un subconjunto finito de este espacio que lo genera; es decir,*

$$\mathbf{E} = \mathrm{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m).$$

Entonces existe $S_1 \subset S$ *que es L.I. y genera a* $\mathbf{E}$*.*

DEMOSTRACIÓN ■ Formemos el conjunto S_1 de la siguiente manera[22]

- $\vec{u}_1 \in S_1$.
- $\vec{u}_2 \in S_1$ si $\vec{u}_2$ no es múltiplo escalar de $\vec{u}_1$.
- En general, $\vec{u}_i \in S_1$ si $\vec{u}_i$ no es combinación lineal de sus predecesores.

Mediante este proceso es claro que los elementos de S_1 son L.I., y ya que cada vector que se excluyó de S para formar S_1 es combinación lineal de sus predecesores, cada uno de ellos se puede escribir como combinación lineal de los elementos de S_1. Así, puesto que todo vector en $\mathbf{E}$ es combinación lineal de los vectores de S, todo vector de $\mathbf{E}$ es combinación lineal de los vectores de S_1; es decir, S_1 genera a $\mathbf{E}$. ■

Corolario 3.2 *Si un espacio vectorial* $\mathbf{E}$ *está generado por los vectores* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ *(no necesariamente L.I.), entonces todo conjunto de vectores linealmente independiente tiene a lo más* m *vectores.*

DEMOSTRACIÓN ■ Por el lema 3.2 podemos suponer, sin pérdida de generalidad, que los vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ son L.I. Si $\vec{v}_i$, $i = 1, \ldots, k$ son vectores L.I., por el corolario 3.1, $k \leq m$. ■

Definición 3.19 *Un espacio vectorial* $\mathbf{E}$ *es finitamente generado si existen* $\vec{u}_1, \ldots, \vec{u}_m \in \mathbf{E}$ *tales que*

$$\mathbf{E} = \mathrm{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m).$$

Corolario 3.3 *Sea* $\mathbf{E}$ *un espacio vectorial finitamente generado, con* $\vec{u}_1, \ldots, \vec{u}_m$ *un conjunto de generadores. Si* $\mathscr{B}$ *es una base de* $\mathbf{E}$*, entonces* $\mathscr{B}$ *tiene a lo más* m *elementos.*

❍ **Nota 3.11** Supongamos que $\mathbf{E}$ es un espacio finitamente generado, con $S = \{\vec{u}_1, \ldots, \vec{u}_m\}$ un conjunto de generadores, entonces, por el lema 3.2, podemos reducir S a un conjunto $\mathscr{B}$ que es linealmente independiente y que también genera a $\mathbf{E}$; es decir, a una base del espacio $\mathbf{E}$.

Teorema 3.16 ***(Igualdad en cardinalidad en bases del mismo espacio)*** *Si* $\mathbf{E}$ *es un espacio vectorial finitamente generado y* $\mathscr{B}_1 = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ *y* $\mathscr{B}_2 = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ *son un par de bases de este espacio, entonces* $n = m$*; es decir, cualquier par de bases en el mismo espacio tienen igual número de elementos.*

[22] Podemos suponer que $\vec{u}_i \neq \vec{0}_{\mathbf{E}}$ para todo i; pues si uno de estos vectores es el neutro aditivo, se puede excluir de S y el conjunto resultante sigue obviamente generando al espacio $\mathbf{E}$.

DEMOSTRACIÓN ■ Por el corolario 3.1, puesto que $\mathscr{B}_2$ es una base y los vectores $\vec{u}_i$ son L.I., se tiene $n \le m$; y por el mismo corolario, dado que los vectores $\vec{v}_i$ son L.I. y $\mathscr{B}_1$ es una base, se deduce $m \le n$; luego $n = m$. ■

Definición 3.20 ***(Dimensión de un espacio vectorial).*** *Sea* $\mathbf{E}$ *un espacio vectorial.*

1. *Si* $\mathbf{E}$ *es un espacio vectorial finitamente generado, se define la* ***dimensión*** *de* $\mathbf{E}$, $\dim(\mathbf{E})$, *como el número de elementos de cualquier base de* $\mathbf{E}$.
2. *Si* $\mathbf{E}$ *no es finitamente generado, entonces se dice que* $\mathbf{E}$ *tiene* ***dimensión infinita*** *y se escribe* $\dim(\mathbf{E}) = \infty$.

► **Ejemplo 3.54** (Dimensión del espacio trivial $\{\vec{0}_{\mathbf{E}}\}$). Sea $\mathbf{E}$ un espacio vectorial, entonces

$$\dim(\{\vec{0}_{\mathbf{E}}\}) = 0;$$

pues el único subconjunto L.I. de este subespacio es el conjunto vacío $\emptyset$ (cfr. ejercicio resuelto 32(a), pág. 191).◄

► **Ejemplo 3.55** (Dimensión de $\mathbb{R}^n$). En el ejemplo 3.50 vimos que

$$\mathscr{B} = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\},$$

donde

$$\vec{e}_i = (\underbrace{0, \ldots, 0, \overset{i}{1}, 0, \ldots, 1}_{n})$$

$i = 1, 2, \ldots, n$ es una base de $\mathbb{R}^n$; por tanto,

$$\dim(\mathbb{R}^n) = n.◄$$

► **Ejemplo 3.56** El espacio de polinomios, $\mathbf{P}$, no es finitamente generado; pues si S es un conjunto finito de generadores, entonces uno de ellos tiene el grado máximo, digamos m. Por tanto, cualquier combinación lineal de elementos de S tiene grado a lo más m y, por ende, un polinomio de grado mayor a m no puede ser combinación lineal de elementos de S. Luego, $\dim(\mathbf{P}) = \infty$ (cfr. ejemplo 3.51).◄

► **Ejemplo 3.57** Probamos, en el ejemplo 3.52, que el conjunto

$$\mathscr{B} = \left\{1, x, x^2, \ldots, x^k\right\}$$

es una base de $\mathbf{P}_k$. Por tanto,

$$\dim(\mathbf{P}_k) = k + 1.◄$$

► **Ejemplo 3.58** En el ejemplo 3.53, mostramos que

$$\mathscr{B} = \{M_{\alpha\beta} \,|\, 1 \le \alpha \le m,\, 1 \le \beta \le n\}$$

es una base del espacio vectorial de las matrices de tamaño $m \times n$. Entonces,

$$\dim(\mathfrak{M}_{m\times n}) = mn.◄$$

▶ **Ejemplo 3.59** En el ejemplo 3.29 vimos que toda matriz simétrica de tamaño 2×2, $\begin{bmatrix} a & c \\ c & d \end{bmatrix}$, se puede escribir como

$$\begin{bmatrix} a & c \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Por otro lado, si

$$\alpha_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \alpha_2\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \alpha_3\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

se desprende que $\alpha_1 = \alpha_2 = \alpha_3 = 0$; i.e., estas matrices son L.I. Luego una base para el espacio, S, de las matrices simétricas de orden 2 es

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

Por tanto,

$$\dim(S) = 3. ◀$$

Extracción de una base en un conjunto generador

En el lema 3.2 probamos que si S es un conjunto finito generador de un espacio vectorial, podemos construir una base del espacio a partir de los elementos de S excluyendo en forma sistemática a los que sean combinaciones lineales de sus predecesores; pues los que queden serán L.I. y claramente aún generan el espacio. Dada la gran importancia que tiene este resultado, volvemos a hacer explícito este método en el siguiente teorema e ilustramos a continuación su uso.

Teorema 3.17 *Sean* $\mathbf{E}$ *un espacio vectorial,* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$*, y supongamos que* $S = \text{gn}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k)$*. Para extraer de* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ *una base de* S*, se excluyen de* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ *aquellos vectores que son combinaciones lineales de sus predecesores.*[23]

▶ **Ejemplo 3.60** Encontrar una base y la dimensión del subespacio de los polinomios

$$S = \text{gn}(2, 1-x, x+1, x^2-1, x^2+1, x^3-2).$$

- Queda 2 como vector inicial para la base.
- Claramente $1-x$ no es múltiplo escalar de 2, por tanto no se excluye.
- Veamos si $1+x$ es combinación lineal de 2 y $1-x$:

$$\begin{aligned} 1+x &= a(2)+b(1-x) \\ &= 2a+b-bx \Rightarrow \\ 2a+b &= 1 \\ -b &= 1; \end{aligned}$$

[23]Por vacuidad, el primer vector no puede ser combinación lineal de sus predecesores y se incluye siempre como vector inicial para la base.

por tanto,

$$b = -1 \text{ y } a = 1.$$

Así que $1+x$ se excluye por ser combinación lineal de sus predecesores.

- x^2-1 no puede ser combinación lineal de 2 y $1-x$, pues toda combinación lineal de estos polinomios tiene grado a lo más uno; luego x^2-1 se queda.
- x^2+1 es combinación lineal de 2, $1-x$ y x^2-1; pues

$$\begin{aligned} x^2+1 &= a(2)+b(1-x)+c(x^2-1) \\ &= 2a+b-c-bx+cx^2 \end{aligned}$$

entraña el sistema

$$\begin{aligned} 2a+b-c &= 1 \\ -b &= 0 \\ c &= 1 \end{aligned}$$

que tiene solución $c=1$, $b=0$ y $a=1$. Por tanto x^2+1 se excluye.

- Finalmente, toda combinación lineal de los vectores que se han conservado tiene grado a lo más 2, por tanto x^3-2 no puede ser combinación lineal de ellos y entonces se conserva.

Luego, la base buscada es

$$\mathscr{B} = \{2, 1-x, x^2-1, x^3-2\}$$

y

$$\dim(S) = 4. ◀$$

En el caso particular de que se quiera extraer una base de $\{\vec{v}_1,\ldots,\vec{v}_k\}$ para el subespacio $S = \text{gn}(\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k)$ de $\mathbb{R}^n$, se puede utilizar herramienta matricial. Sea A la matriz que tiene por columnas a los vectores $\vec{v}_i$ y H la forma escalonada reducida de A. Si $\vec{b} \in S$, entonces $\vec{b}$ es combinación lineal de los vectores columna de A que corresponden a columnas con pivote en H; pues para los sistemas equivalentes $[A\,|\,\vec{b}] \sim [H\,|\,\vec{c}]$ estas columnas son precisamente donde hay variables ligadas y en las restantes columnas hay variables libres; a cada una de estas últimas se le puede asignar cualquier valor, en particular valores nulos a todas. Así S está generado por los vectores columna de A que corresponden a columnas con pivote en cualquier matriz escalonada equivalente a la matriz A.[24] Ahora pensemos en la matriz A_1 que se obtiene de la matriz A excluyendo los vectores que no correspondieron a columnas con pivote en la forma escalonada reducida; digamos que A_1 tiene por columnas a $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_m$. Entonces, toda columna de la forma escalonada reducida de A_1 tiene pivote; luego el sistema $A_1\vec{x}=\vec{0}$, $\vec{x}\in\mathbb{R}^m$, tiene solución única, y por tanto las columnas de A_1 son L.I. Con lo precedente, hemos hecho plausible la demostración del siguiente teorema, que es un método para extraer una base de un subespacio generado por vectores de $\mathbb{R}^n$.

[24] Cfr. teorema 3.10 y la discusión que lo precede en la página 150.

Teorema 3.18 *Sean* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in \mathbb{R}^n$ *y* $S = \operatorname{gn}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k)$. *Entonces, para obtener una base de* S *a partir de los vectores* $\vec{v}_i$, *se procede de la manera siguiente:*

1. *Se forma la matriz* A *que tiene por columnas los vectores* $\vec{v}_i$ *y se lleva a forma escalonada.*
2. *Los vectores columna de* A *que correspondan a columnas con pivote en esa forma escalonada formarán una base de* S.

► **Ejemplo 3.61** Extraer una base y encontrar la dimensión del subespacio S generado por los vectores $(1,-1,0,2,1)$, $(2,1,-2,0,0)$, $(0,-3,2,4,2)$, $(2,4,1,0,1)$, $(3,3,-4,-2,-1)$, $(5,7,-3,-2,0)$.◄

Solución Procedamos como establece el teorema 3.18:

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 3 & 5 \\ -1 & 1 & -3 & 4 & 3 & 7 \\ 0 & -2 & 2 & 1 & -4 & -3 \\ 2 & 0 & 4 & 0 & -2 & -2 \\ 1 & 0 & 2 & 1 & -1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 2 & 3 & 5 \\ 0 & 3 & -3 & 6 & 6 & 12 \\ 0 & -2 & 2 & 1 & -4 & -3 \\ 0 & -4 & 4 & -4 & -8 & -12 \\ 0 & -2 & 2 & -1 & -4 & -5 \end{bmatrix}$$
$$\to \begin{bmatrix} 1 & 2 & 0 & 2 & 3 & 5 \\ 0 & 1 & -1 & 2 & 2 & 4 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

La base buscada se forma tomando los vectores columna de la matriz inicial que correspondan a columnas que tengan pivote en la forma escalonada

$$\mathscr{B} = \{(1,-1,0,2,1),(2,1,-2,0,0),(2,4,1,0,1)\}$$

y $\dim(S) = 3$. ✔

El siguiente teorema nos auxilia para, en el caso de saber la dimensión de un espacio, digamos n, concluir que un conjunto de n generadores son L.I., y por ende forman una base, o que n vectores L.I. generan al espacio; y así evitar probar la independencia lineal de los vectores o que generen al espacio.

Teorema 3.19 *Si* $\dim(\mathbf{E}) = n$, *las siguientes condiciones son equivalentes:*

1. $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ *son L.I.*
2. $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ *generan a* $\mathbf{E}$.
3. $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ *es una base de* $\mathbf{E}$.

DEMOSTRACIÓN ■ Sea $\mathscr{B}$ una base de $\mathbf{E}$, entonces $\mathscr{B}$ tiene n elementos pues $\dim(\mathbf{E}) = n$. Mostremos (1) $\Rightarrow$ (2) $\Rightarrow$ (3), (3) $\Rightarrow$ (1) y tendremos la equivalencia de las tres condiciones.

(1) $\Rightarrow$ (2). Supongamos que $\vec{w} \in \mathbf{E}$, entonces, por el teorema 3.15, los vectores $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n, \vec{w}\}$ son L.D. Entonces existen escalares α_i, $i = 1, 2, \ldots, n$, y $\beta \in \mathbb{R}$, no todos cero, tales que

$$\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_n\vec{u}_n + \beta\vec{w} = \vec{0}_{\mathbf{E}}.$$

Si $\beta = 0$, entonces algún α_i debe ser distinto de cero; pero esto es imposible ya que los vectores $\vec{u}_i$ son L.I.; por tanto $\beta \neq 0$. Luego $\vec{w}$ es combinación lineal de los vectores $\vec{u}_i$. Dado que $\vec{w}$ es arbitrario, hemos probado que los vectores $\vec{u}_i$ generan a **E**.

(2) $\Rightarrow$ (3). Si los vectores $\vec{u}_i$ son L.D., entonces, por el teorema 3.17, podemos reducir este conjunto a uno L.I. que genere a **E** con m elementos y $m < n$. Pero esto implicaría que hay una base de **E** con un número de elementos distinto de n, lo cual es una contradicción al teorema 3.16. Por tanto, los vectores $\vec{u}_i$ deben ser L.I. Luego forman una base de **E**.

(3) $\Rightarrow$ (1). Es clara. ■

▶ **Ejemplo 3.62** Sabemos que $\mathbf{P}_2$ tiene dimensión 3. Si $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ son tales que

$$\alpha_1 (x-2) + \alpha_2 (3-x) + \alpha_3 \left(x^2 - x\right) = 0,$$

entonces

$$3\alpha_2 - 2\alpha_1 + (\alpha_1 - \alpha_2 - \alpha_3)x + \alpha_3 x^2 = 0.$$

Por la definición de igualdad de polinomios se tiene

$$\begin{aligned} 3\alpha_2 - 2\alpha_1 &= 0 \\ \alpha_1 - \alpha_2 - \alpha_3 &= 0 \\ \alpha_3 &= 0, \end{aligned}$$

que implica $\alpha_3 = \alpha_2 = \alpha_1 = 0$. Luego los tres polinomios $x-2$, $3-x$, x^2-x son L.I. en $\mathbf{P}_2$ y por tanto $\left\{x-2, 3-x, x^2-x\right\}$ es una base de $\mathbf{P}_2$.◀

Compleción de un conjunto L.I. a una base

Podemos utilizar el teorema 3.17 para completar un conjunto L.I. a una base $\mathscr{B}$ del espacio mediante el siguiente procedimiento:

Sea $\mathscr{B}_1 = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ un conjunto L.I. de un espacio **E** que tiene dimensión finita n.

1. Se forma el conjunto

$$M = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\},$$

   donde $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ es una base conocida del espacio **E**.
2. Se extrae de M una base $\mathscr{B}$ por medio del procedimiento dado en el teorema 3.17.
3. El conjunto $\mathscr{B}$ encontrado en el inciso anterior será una base de **E**.

❍ **Nota 3.12** El procedimiento anterior se justifica porque $\mathbf{E} = \mathrm{gn}(M)$ y, al aplicar el teorema 3.17 a este conjunto generador, los vectores $\vec{v}_i$, $i = 1, 2, \ldots, k$, permanecen ya que por hipótesis son L.I.; luego la base $\mathscr{B}$ del espacio **E** contiene a estos vectores.

▶ **Ejemplo 3.63** Completar el conjunto L.I. $\{2x-4, x^2+1\}$ a una base de $\mathbf{P}_3$.◀

Solución Utilizamos como base conocida de $\mathbf{P}_3$ a $\{1,x,x^2,x^3\}$ y extraemos de $\{2x-4,x^2+1,1,x,x^2,x^3\}$ una base:

- $2x-4$ y x^2+1 se quedan porque son L.I.
- $1=a(2x-4)+b(x^2+1)$ implica $bx^2+2ax+b-4a=1$; por tanto,

$$\begin{aligned} b &= 0 \\ 2a &= 0 \\ b-4a &= 1, \end{aligned}$$

  sistema que es inconsistente; luego 1 no es combinación lineal de sus predecesores y se queda.
- $x=a(2x-4)+b(x^2+1)+c$ implica $x=bx^2+2ax+b+c-4a$; por tanto,

$$\begin{aligned} b &= 0 \\ 2a &= 1 \\ b+c-4a &= 0, \end{aligned}$$

  sistema consistente con $b=0$, $a=1/2$ y $c=2$; luego x es combinación lineal de sus predecesores y se excluye.
- Evidentemente $x^2=0(2x-4)+1(x^2+1)+(-1)1$; así que x^2 es combinación lineal de sus predecesores y se excluye.
- Claramente x^3 no es combinación lineal de sus predecesores, así que se incluye.

La base buscada es

$$\mathscr{B}=\{2x-4,x^2+1,1,x^3\}. \quad ✔$$

▶ **Ejemplo 3.64** Completar el conjunto L.I. $\{(1,2,0,1,3),(-2,-4,1,-2,-6)\}$ a una base de $\mathbb{R}^5$.◀

Solución Utilizamos la base canónica de $\mathbb{R}^5$, la adjuntamos a este conjunto

$$\begin{aligned} M &= \{(1,2,0,1,3),(-2,-4,1,-2,-6),(1,0,0,0,0),(0,1,0,0,0), \\ &\quad (0,0,1,0,0),(0,0,0,1,0),(0,0,0,0,1)\}, \end{aligned}$$

y extraemos una base del mismo utilizando el teorema 3.18:

$$\begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 2 & -4 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & -2 & 0 & 0 & 0 & 1 & 0 \\ 3 & -6 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 & 1 \end{bmatrix}.$$

La base buscada es

$$\{(1,2,0,1,3),(-2,-4,1,-2,-6),(1,0,0,0,0,0),(0,1,0,0,0),(0,0,0,1,0)\}. \quad ✔$$

3.4.3 Rango de una matriz

Recordemos que si $A \in \mathfrak{M}_{m\times n}$, el subespacio de $\mathbb{R}^m$ generado por las columnas de A es el espacio columna de A, y el subespacio de $\mathbb{R}^n$ generado por las filas de A es el espacio fila de A y se les denota, respectivamente, por $E_c(A)$ y $E_f(A)$ (cfr. definición 3.13, pág. 146); y que el espacio nulo de una matriz se define como el subespacio de $\mathbb{R}^n$ de las soluciones del sistema homogéneo $A\vec{x} = \vec{0}$ (cfr. ejemplo 3.20, pág. 141).

Definición 3.21 *Si $A \in \mathfrak{M}_{m\times n}$ se definen:*

1. *Rango fila de A:* $\mathrm{R_f}(A) = \dim(E_f(A))$.
2. *Rango columna de A:* $\mathrm{R_c}(A) = \dim(E_c(A))$.
3. *Nulidad de A,* $\mathrm{Nul}(A)$*, como la dimensión del espacio nulo de A.*

Supongamos que H es la forma escalonada reducida de una matriz $A \in \mathfrak{M}_{m\times n}$. Como vimos anteriormente, las columnas de A que corresponden a columnas con pivote de H generan al espacio columna y son linealmente independientes; además, hay tantas columnas en H con pivote como filas no nulas en H. Así que el rango columna de A es el número de filas no nulas de cualquier forma escalonada equivalente a la matriz A. Por otra parte, las filas de H se obtienen a través de sucesiones de operaciones de renglón iniciando con filas de la matriz A; luego toda fila de H es elemento del espacio fila de A. Además, las filas no nulas de H son linealmente independientes; pues en caso contrario, si una fila no nula de H es combinación lineal de otras filas, se pueden aplicar operaciones de renglón para transformar esta fila en una fila nula; lo cual no es posible, pues H es la forma escalonada reducida de A (recuerde que la forma escalonada reducida de una matriz es única). Entonces tenemos que las filas no nulas de cualquier matriz en forma escalonada equivalente a la matriz A forman una base para el espacio fila; luego el rango fila de A es el número de filas no nulas de cualquier forma escalonada equivalente a la matriz A. Finalmente, si en H existen vectores columna sin pivote, para fijar ideas pensemos que son 2 y que se encuentran en las dos últimas columnas, entonces hay dos variables libres r y s, y las soluciones del sistema homogéneo son de la forma

$$\begin{bmatrix} a_1r+b_1s \\ a_2r+b_2s \\ \vdots \\ a_{n-2}r+b_{n-2}s \\ s \\ r \end{bmatrix} = r\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-2} \\ 0 \\ 1 \end{bmatrix} + s\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-2} \\ 1 \\ 0 \end{bmatrix}.$$

Así, toda solución es combinación lineal de los vectores $\vec{u}_1 = (a_1,\ldots,a_{n-2},0,1)$ y $\vec{u}_2 = (b_1,\ldots,b_{n-2},1,0)$. Si estos últimos vectores fueran L.D., entonces el sistema tendría sólo una variable libre, lo cual es imposible. Con esto hemos hecho plausible la demostración del siguiente teorema.

Teorema 3.20 *Sean $A \in \mathfrak{M}_{m\times n}$ y H una forma escalonada equivalente, entonces:*

1. *Una base del espacio fila de A la forman las filas no nulas de H.*
2. *Una base del espacio columna de A la forman las columnas de A que corresponden a columnas con pivote de H.*
3. *El rango fila de A es el número de filas no nulas de H.*
4. *El rango columna de A es el número de columnas con pivote de H.*
5. $\mathrm{R_f}(A) = \mathrm{R_c}(A)$.
6. *La nulidad de A es el número de columnas sin pivote de H.*

▶ **Ejemplo 3.65** Hallar:

1. Una base de $E_f(A)$ y $\mathrm{R_f}(A)$.
2. Una base de $E_c(A)$ y $\mathrm{R_c}(A)$.
3. Una base de espacio nulo de A y $\mathrm{Nul}(A)$.

Si A es la matriz

$$\begin{bmatrix} 1 & 2 & 0 & -1 & 1 \\ 1 & 3 & 1 & 1 & -1 \\ 2 & 5 & 1 & 0 & 0 \\ 3 & 6 & 0 & 0 & -6 \\ 1 & 5 & 3 & 5 & -5 \end{bmatrix}. ◀$$

Solución Llevemos la matriz a forma escalonada.

$$A \to \begin{bmatrix} 1 & 2 & 0 & -1 & 1 \\ 0 & 1 & 1 & 2 & -2 \\ 0 & 1 & 1 & 2 & -2 \\ 0 & 0 & 0 & 3 & -9 \\ 0 & 3 & 3 & 6 & -6 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & -1 & 1 \\ 0 & 1 & 1 & 2 & -2 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Entonces:

1. Las filas distintas de cero de la forma escalonada constituyen una base para el espacio fila, luego

$$\mathscr{B} = \{(1,2,0,-1,1),(0,1,1,2,-2),(0,0,0,1,-3)\},$$

   y $\mathrm{R_f}(A) = 3$.

2. Las columnas de A correspondientes a columnas con pivote en la forma escalonada forman una base del espacio columna, así

$$\mathscr{B} = \{(1,1,2,3,1),(2,3,5,6,5),(1-1,1,0,0,5)\},$$

   y $\mathrm{R_c}(A) = 3$.

3. La solución al sistema homogéneo es:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2s+10r \\ -s-4r \\ s \\ 3r \\ r \end{bmatrix} = r\begin{bmatrix} 10 \\ -4 \\ 0 \\ 3 \\ 1 \end{bmatrix} + s\begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Entonces, una base para el espacio nulo es

$$\mathscr{B} = \{(10,-4,0,3,1),(2,-1,1,0,0)\},$$

y $\mathrm{Nul}(A) = 2$. ✔

Definición 3.22 *Si $A \in \mathfrak{M}_{m\times n}$, el rango de A se define como*

$$\mathrm{Rang}(A) = \mathrm{R_f}(A) = \mathrm{R_c}(A).$$

► **Ejemplo 3.66** Si A es la matriz del ejemplo precedente,

$$\mathrm{Rang}(A) = 3. ◄$$

Los dos siguientes teoremas son conclusiones inmediatas de lo precedente y su demostración se deja como ejercicio al lector.

Teorema 3.21 *Si $A \in \mathfrak{M}_{m\times n}$, entonces $\mathrm{Rang}(A) + \mathrm{Nul}(A) = n$. Equivalentemente para el sistema $A\vec{x} = \vec{0}$:*

1. $\mathrm{Rang}(A)$ = número de variables ligadas.
2. $\mathrm{Nul}(A)$ = número de variables libres.
3. $\mathrm{Rang}(A) + \mathrm{Nul}(A) = n$.

Teorema 3.22 *Si $A \in \mathfrak{M}_{n\times n}$, A es invertible si y sólo si $\mathrm{Rang}(A) = n$.*

Para finalizar este apartado enunciamos y probamos el siguiente teorema que será útil más adelante.

Teorema 3.23 *Si $A \in \mathfrak{M}_{m\times n}$, entonces la matriz simétrica A^tA y la matriz A tienen el mismo rango; esto es,*

$$\mathrm{Rang}(A^tA) = \mathrm{Rang}(A).$$

DEMOSTRACIÓN ■ Mostremos que los sistemas $A\vec{x} = \vec{0}$ y $(A^tA)\vec{x} = \vec{0}$ tienen las mismas soluciones.

1. Si $\vec{u} \in \mathbb{R}^n$ es solución de $A\vec{x} = \vec{0}$, entonces $A\vec{u} = \vec{0}$; luego

$$A^t\,(A\vec{u}) = \vec{0};$$

por tanto,

$$(A^tA)\,\vec{u} = \vec{0}.$$

2. Supongamos ahora que $\vec{v} \in \mathbb{R}^n$ es solución de $(A^tA)\,\vec{x} = \vec{0}$. Entonces

$$(A^tA)\,\vec{v} = \vec{0};$$

por tanto,

$$(\vec{v})^t\,(A^tA)\,\vec{v} = 0;$$

esto es,

$$(\vec{v}^{\,t}A^t)A\vec{v} = 0;$$

es decir,

$$(A\vec{v})^t\,(A\vec{v}) = 0;$$

por ende,

$$\|A\vec{v}\|^2 = 0,$$

por lo que

$$A\vec{v} = \vec{0}.$$

Luego, de 1 y 2, los sistemas $A\vec{x} = \vec{0}$ y $(A^tA)\vec{x} = \vec{0}$ tienen las mismas soluciones; por tanto,

$$\mathrm{Nul}(A) = \mathrm{Nul}(A^tA),$$

y dado que A tiene tamaño $m \times n$ y A^tA tiene tamaño $n \times n$, del teorema 3.21 inciso 3 se desprende que

$$\mathrm{Rang}(A) = \mathrm{Rang}(A^tA). \quad ■$$

► **Ejemplo 3.67** Sea

$$A = \begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & -1 & 0 & 1 \\ -2 & 1 & 1 & 0 \end{bmatrix}.$$

Llevemos A a forma escalonada

$$A = \begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & -1 & 0 & 1 \\ -2 & 1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -4 & -1 \\ 0 & -1 & 5 & 2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -4 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Entonces

$$\operatorname{Rang}(A) = 3.$$

Por otro lado,

$$A^tA = \begin{bmatrix} 1 & 2 & -2 \\ -1 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & -1 & 0 & 1 \\ -2 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 9 & -5 & 0 & 3 \\ -5 & 3 & -1 & -2 \\ 0 & -1 & 5 & 2 \\ 3 & -2 & 2 & 2 \end{bmatrix}.$$

Llevemos esta última matriz a forma escalonada

$$\begin{bmatrix} 9 & -5 & 0 & 3 \\ -5 & 3 & -1 & -2 \\ 0 & -1 & 5 & 2 \\ 3 & -2 & 2 & 2 \end{bmatrix} \sim \begin{bmatrix} 9 & -5 & 0 & 3 \\ -5 & 3 & -1 & -2 \\ 0 & -1 & 5 & 2 \\ 1 & -1 & 3 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 3 & 2 \\ -5 & 3 & -1 & -2 \\ 0 & -1 & 5 & 2 \\ 9 & -5 & 0 & 3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 14 & 8 \\ 0 & -1 & 5 & 2 \\ 0 & 4 & -27 & -15 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 3 & 2 \\ 0 & 1 & -7 & -4 \\ 0 & -1 & 5 & 2 \\ 0 & 4 & -27 & -15 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -1 & 3 & 2 \\ 0 & 1 & -7 & -4 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 3 & 2 \\ 0 & 1 & -7 & -4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

De donde

$$\operatorname{Rang}(A^tA) = 3. \blacktriangleleft$$

3.5 Espacios vectoriales sobre los números complejos

Es posible considerar espacios vectoriales sobre el campo de los números complejos.[25] Esto significa un conjunto **E** en el que se ha definido una operación suma entre sus elementos (vectores) y una operación producto de un escalar por un vector, siendo los escalares números complejos, que cumplen los diez axiomas de espacio vectorial de la definición 3.6 (cfr. pág. 131).[26] De ser así, se tiene un **espacio vectorial sobre el campo de los números complejos** o simplemente un **espacio complejo**. Todo lo desarrollado en este libro relativo a espacios vectoriales reales (los escalares subyacentes son

[25]Recomendamos al lector consultar B.1 del apéndice B y los apartados 1.1.5, 1.2.8, 2.1.5 y 2.2.5.

[26]Obviamente reemplazando el símbolo $\mathbb{R}$ por el símbolo $\mathbb{C}$ donde corresponda.

números reales) tiene su símil para espacios sobre el campo de los números complejos, y toda la teoría sigue siendo válida para este tipo de espacios pues depende, en cuanto a la multiplicación de vectores por escalares, únicamente de las propiedades de campo de los números reales, que son las mismas que las de los números complejos, y no de propiedades particulares de los números reales como el orden (propiedad totalmente ausente en $\mathbb{C}$). Es decir, los conceptos de subespacio vectorial, combinaciones lineales, subespacios generados, independencia lineal, bases, dimensión, etc., siguen siendo los mismos; excepto que el conjunto de escalares es el campo $\mathbb{C}$.

► **Ejemplo 3.68** Consideremos $\mathfrak{M}_{m\times n}(\mathbb{C})$ el conjunto de matrices de tamaño $m \times n$ cuyas componentes son números complejos con la suma usual de matrices (sumando en forma compleja componente a componente) y multiplicación de un escalar por una matriz; pero esta vez los escalares son también números complejos. Es fácil verificar que $\mathfrak{M}_{m\times n}(\mathbb{C})$ es un espacio vectorial sobre $\mathbb{C}$ y que $\dim(\mathfrak{M}_{m\times n}(\mathbb{C})) = mn$.◄

► **Ejemplo 3.69** Otro ejemplo es $\mathbb{C}^n$, el conjunto de n-adas ordenadas $(z_1,\ldots,z_n)$ donde $z_i \in \mathbb{C}$ para cada $i = 1,\ldots,n$; con las operaciones

$$\begin{aligned}(z_1,z_2,\ldots,z_n)+(w_1,w_2,\ldots,w_n) &= (z_1+w_1,z_2+w_2,\ldots,z_n+w_n),\\ \alpha(z_1,z_2,\ldots,z_n) &= (\alpha z_1,\alpha z_2,\ldots,\alpha z_n),\end{aligned}$$

donde z_j+w_j y αz_j son las operaciones de suma y multiplicación de números complejos, es fácil probar con estas operaciones que $\mathbb{C}^n$ es un espacio vectorial sobre el campo $\mathbb{C}$, y que

$$\dim(\mathbb{C}^n) = n$$

(una base para este espacio es también la base canónica de $\mathbb{R}^n$).◄

► **Ejemplo 3.70** Sea el conjunto $\mathbf{E} = \mathbb{C}^n$ considerado en el ejemplo precedente con la suma entre sus elementos ahí definida, pero con el producto de un escalar por un vector restringido a los números reales. Es claro que $\mathbf{E}$ es un espacio vectorial real. Sean los vectores

$$\begin{aligned}\vec{e}_j &= (\underbrace{0,\ldots,0,\overset{j}{1},0,\ldots,0}_{n}) \text{ para } j=1,\ldots,n, \quad \text{y}\\ \vec{f}_j &= (\underbrace{0,\ldots,0,\overset{j}{i},0,\ldots,0}_{n}) \text{ para } j=1,\ldots,n.\end{aligned}$$

Claramente los $2n$ vectores $\vec{e}_j,\vec{f}_j$, $j=1,\ldots,n$, son L.I., y si

$$\vec{w} = (a_1+b_1 i,\ldots,a_n+b_n i),$$

entonces

$$\vec{w} = \sum_{j=1}^{n} a_j\vec{e}_j + \sum_{j=1}^{n} b_j\vec{f}_j;$$

de donde se desprende que

$$\dim(\mathbf{E}) = 2n.◄$$

❍ **Nota 3.13** Los espacios sobre el campo de los números complejos son de gran importancia en álgebra lineal y están involucrados directa o indirectamente en una gran variedad de aplicaciones de esta rama de las matemáticas. Se pudo considerar este tipo de espacios desde el inicio, pero por cuestiones que tienen que ver con aspectos psicológicos, debido a la poca familiaridad que por lo general tienen los lectores con el campo $\mathbb{C}$, no se hizo y se han incluido secciones donde se ha intentado introducir estos conceptos gradualmente con el fin de que el lector se vaya adaptando de esta misma forma a espacios sobre $\mathbb{C}$. Es hasta el final del capítulo 5 que se verá la importancia del uso de escalares complejos y de espacios sobre este campo. Por esta razón, de aquí en adelante supondremos que todos los espacios vectoriales con los que se trabaje serán espacios vectoriales reales, a menos que se indique lo contrario; en cuyo caso se hará notar que se tiene un espacio complejo.

3.6 Ejercicios resueltos y ejercicios propuestos

3.6.1 Ejercicios resueltos

Geometría de los espacios $\mathbb{R}^n$

En los ejercicios 1 a 8, $\vec{u} = (-1,3,4,-2)$, $\vec{v} = (-2,5,3,2)$ y $\vec{w} = (1,0,4,-2)$.

1 Calcular $3\vec{u} - 4\vec{v}$.

Solución

$$\begin{aligned} 3\vec{u} - 4\vec{v} &= 3(-1,3,4,-2) - 4(-2,5,3,2) \\ &= (-3,9,12,-6) - (-8,20,12,8) \\ &= (5,-11,0,-14). \quad ✔ \end{aligned}$$

2 Encontrar $\vec{u} + \frac{1}{2}\vec{w}$.

Solución

$$\begin{aligned} \vec{u} + \tfrac{1}{2}\vec{w} &= (-1,3,4,-2) + \tfrac{1}{2}(1,0,4,-2) \\ &= (-1,3,4,-2) + \left(\tfrac{1}{2},0,2,-1\right) \\ &= \left(-\tfrac{1}{2},3,6,-3\right). \quad ✔ \end{aligned}$$

3 Hallar $\|\vec{v}\|$.

Solución

$$\|\vec{v}\| = \sqrt{(-2)^2 + (5)^2 + (3)^2 + (2)^2} = \sqrt{42}. \quad ✔$$

4 Calcular $\mathrm{d}(\vec{u},\vec{w})$.

Solución

$$\begin{aligned} \mathrm{d}(\vec{u},\vec{w}) &= \|\vec{u} - \vec{w}\| \\ &= \|(-1,3,4,-2) - (1,0,4,-2)\| \\ &= \|(-2,3,0,0)\| \\ &= \sqrt{13}. \quad ✔ \end{aligned}$$

5 Hallar $4\vec{u}-2\vec{v}-\vec{w}$.

Solución

$$\begin{aligned} 4\vec{u}-2\vec{v}-\vec{w} &= 4(-1,3,4,-2)-2(-2,5,3,2)-(1,0,4,-2)\\ &= (-4,12,16,-8)-(-4,10,6,4)-(1,0,4,-2)\\ &= (0,2,10,-12)-(1,0,4,-2)\\ &= (-1,2,6,-10)\,. \quad ✔ \end{aligned}$$

6 Encontrar un vector unitario (con norma 1) paralelo y con la misma dirección que el vector $\vec{v}$.

Solución Por el ejercicio 3, $\|\vec{v}\| = \sqrt{42}$; así que un vector unitario paralelo en la misma dirección está dado por

$$\begin{aligned} \vec{a} &= \tfrac{1}{\|\vec{v}\|} = \tfrac{1}{\sqrt{42}}(-2,5,3,2)\\ &= \left(-\tfrac{1}{21}\sqrt{42}, \tfrac{5}{42}\sqrt{42}, \tfrac{1}{14}\sqrt{42}, \tfrac{1}{21}\sqrt{42}\right). \quad ✔ \end{aligned}$$

7 Calcular $\vec{u}\cdot(\vec{v}+\vec{w})$.

Solución

$$\begin{aligned} \vec{u}\cdot(\vec{v}+\vec{w}) &= (-1,3,4,-2)\cdot((-2,5,3,2)+(1,0,4,-2))\\ &= (-1,3,4,-2)\cdot(-1,5,7,0)\\ &= 44. \quad ✔ \end{aligned}$$

8 Encontrar el valor de b tal que $(-1,b,1,-2)\perp\vec{u}$; es decir, para que estos vectores sean ortogonales.

Solución Para que $(-1,b,1,-2)\perp\vec{u}$ es necesario y suficiente que $(-1,b,1,-2)\cdot\vec{u}=0$. Entonces,

$$\begin{aligned} (-1,b,1,-2)\cdot(-1,3,4,-2) &= 0 \Rightarrow\\ 9+3b &= 0 \Rightarrow\\ b &= -3. \quad ✔ \end{aligned}$$

9 Calcular el área del paralelogramo generado por los vectores $(3,2)$ y $(8,0)$ utilizando la interpretación geométrica del determinante del apartado 3.1.2 (cfr. pág. 117). Comprobar que el resultado es correcto mediante la fórmula usual para el área de un paralelogramo.

Solución Sea $A = \begin{bmatrix} 3 & 8 \\ 2 & 0 \end{bmatrix}$, entonces el área a del paralelogramo es

$$\begin{aligned} a &= |\det(A)|\\ &= 16\ \text{u}^2. \end{aligned}$$

El área de un paralelogramo es el producto de la base b por la altura h; en este caso $b=8$ y $h=2$; por tanto, $a=bh=16\text{u}^2$. ✔

En los ejercicios 10 y 11, los vectores de la base canónica de $\mathbb{R}^3$ $\vec{e}_1=(1,0,0)$, $\vec{e}_2=(0,1,0)$ y $\vec{e}_3=(0,0,1)$ serán representados de la manera tradicional por $\widehat{\imath}$, $\widehat{\jmath}$ y $\widehat{k}$, respectivamente, y $\vec{u}=(a_1,a_2,a_3)$, $\vec{v}=(b_1,b_2,b_3)$ son cualquier par de vectores *no paralelos de* $\mathbb{R}^3$.

10 Encontrar el conjunto de vectores $\vec{w}=(c_1,c_2,c_3)\in\mathbb{R}^3$ que son perpendiculares tanto a $\vec{u}$ como a $\vec{v}$.

Solución Si $a_1=0$ y $b_1=0$, entonces el conjunto buscado consiste en los vectores de la forma $\alpha\widehat{\imath}$. Se puede suponer entonces, sin perder generalidad, que $a_1\neq 0$. Se debe tener entonces $\vec{u}\cdot\vec{w}=0=\vec{v}\cdot\vec{w}$; esto es

$$
\begin{aligned}
a_1c_1 + a_2c_2 + a_3c_3 &= 0,\\
b_1c_1 + b_2c_2 + b_3c_3 &= 0.
\end{aligned}
$$

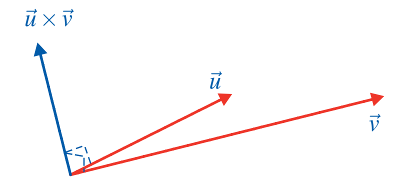
Y ya que

$$
\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \sim \begin{bmatrix} a_1 & a_2 & a_3 \\ 0 & a_1b_2 - a_2b_1 & b_3a_1 - a_3b_1 \end{bmatrix},
$$

y $a_1b_2 - a_2b_1 \neq 0$, porque $\vec{u}$ y $\vec{v}$ no son paralelos, se tiene que el conjunto de vectores ortogonales a $\vec{u}$ y $\vec{v}$ está dado por:

$$
\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{a_1}\left(a_2\frac{a_3b_1 - a_1b_3}{a_1b_2 - a_2b_1}r + a_3r\right) \\ \frac{a_3b_1 - a_1b_3}{a_1b_2 - a_2b_1}r \\ r \end{bmatrix}, \quad r \in \mathbb{R}. \checkmark
$$

11 (**Producto cruz de vectores o producto vectorial en** $\mathbb{R}^3$). Elíjase $r = a_1b_2 - a_2b_1$ en el inciso precedente para calcular el vector resultante $\vec{p}$ que es ortogonal a $\vec{u}$ y $\vec{v}$. Encontrar una fórmula que involucre al determinante de una matriz adecuada para calcular este vector. Al vector $\vec{p}$ se le llama producto cruz o producto vectorial de los vectores $\vec{u}$ y $\vec{v}$ y se acostumbra emplear la notación $\vec{p} = \vec{u} \times \vec{v}$.



Solución En este caso

$$
c_3 = a_1b_2 - a_2b_1 = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix},
$$

$$
\begin{aligned}
c_2 &= \frac{a_3b_1 - a_1b_3}{a_1b_2 - a_2b_1}r \\
&= a_3b_1 - a_1b_3 \\
&= \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix},
\end{aligned}
$$

y

$$
\begin{aligned}
c_1 &= -\frac{1}{a_1}\left(a_2\frac{a_3b_1 - a_1b_3}{a_1b_2 - a_2b_1}r + a_3r\right) \\
&= -\frac{1}{a_1}(a_2(a_3b_1 - a_1b_3) + a_3(a_1b_2 - a_2b_1)) \\
&= \frac{1}{a_1}(a_2a_1b_3 - a_3a_1b_2) \\
&= a_2b_3 - a_3b_2 \\
&= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}.
\end{aligned}
$$

Luego,

$$\begin{aligned}\vec{p} &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\hat{\imath} + \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix}\hat{\jmath} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\hat{k} \\ &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\hat{\imath} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}\hat{\jmath} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\hat{k} \\ &= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix},\end{aligned}$$

al desarrollar el último determinante por cofactores en la primera fila. ✔

12 Encontrar el producto vectorial de $\vec{u} = (-1, 2, -1)$ y $\vec{v} = (2, -1, 3)$.

Solución

$$\begin{aligned}\vec{u}\times\vec{v} &= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ -1 & 2 & -1 \\ 2 & -1 & 3 \end{vmatrix} \\ &= \begin{vmatrix} 2 & -1 \\ -1 & 3 \end{vmatrix}\hat{\imath} - \begin{vmatrix} -1 & -1 \\ 2 & 3 \end{vmatrix}\hat{\jmath} + \begin{vmatrix} -1 & 2 \\ 2 & -1 \end{vmatrix}\hat{k} \\ &= 5\hat{\imath} + \hat{\jmath} - 3\hat{k} \\ &= (5, 1, -3).\end{aligned}$$ ✔

13 Calcular la magnitud del producto vectorial de cualquier par de vectores $\vec{u} = (a_1, a_2, a_3)$ y $\vec{v} = (b_1, b_2, b_3)$ de $\mathbb{R}^3$ en términos del ángulo θ entre ellos, e interpretar geométricamente el significado de esta magnitud.

Solución En el apartado 3.1.2 (cfr. pág. 117) se dedujo que el área del paralelogramo generado por estos dos vectores (cfr. figura 3-9) está dada por

$$S = \|\vec{u}\|\,\|\vec{v}\| \operatorname{sen}\theta,$$

donde θ es el ángulo entre ellos; también se concluyó ahí que

$$\begin{aligned}S^2 &= \|\vec{u}\|^2\,\|\vec{v}\|^2 - (\vec{u}\cdot\vec{v})^2 \\ &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2 + b_2 + a_3b_3)^2.\end{aligned}$$

Por otra parte,

$$\begin{aligned}\|\vec{u}\times\vec{v}\|^2 &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2 \\ &= (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2 \\ &= a_2^2b_3^2 - 2a_2b_3a_3b_2 + a_3^2b_2^2 + a_3^2b_1^2 - 2a_3b_1a_1b_3 + a_1^2b_3^2 + a_1^2b_2^2 - 2a_1b_2a_2b_1 + a_2^2b_1^2 \\ &= a_2^2b_3^2 + a_3^2b_2^2 + a_3^2b_1^2 + a_1^2b_3^2 + a_1^2b_2^2 + a_2^2b_1^2 + a_1^2b_1^2 + a_2^2b_2^2 + a_3^2b_3^2 - a_1^2b_1^2 - a_2^2b_2^2 - a_3^2b_3^2 \\ &\quad -2a_2b_3a_3b_2 - 2a_3b_1a_1b_3 - 2a_1b_2a_2b_1 \\ &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - \left((a_1b_1 + a_2b_2)^2 + 2(a_1b_1a_2b_2)a_3b_3 + a_3^2b_3^2\right) \\ &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2 \\ &= S^2.\end{aligned}$$

De donde

$$S = \|\vec{u} \times \vec{v}\|$$

y

$$\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| \, |\operatorname{sen}\theta| \, . \quad ✔$$

14 Demostrar las propiedades del espacio vectorial de $\mathbb{R}^n$ (cfr. pág. 122).

DEMOSTRACIÓN ■ Sean $\vec{u} = (x_1, \ldots, x_n)$, $\vec{v} = (y_1, \ldots, y_n)$, $\vec{w} = (z_1, \ldots, z_n)$ vectores arbitrarios de $\mathbb{R}^n$ y λ, β cualquier par de números reales. Entonces:

1. Dado que la suma de números reales es también un número real,

$$\vec{u} + \vec{v} = (x_1 + y_1, \ldots, x_n + y_n) \in \mathbb{R}^n.$$

2. Puesto que la suma de números reales es asociativa,

$$\begin{aligned} \vec{u} + (\vec{v} + \vec{w}) &= (x_1, \ldots, x_n) + ((y_1, \ldots, y_n) + (z_1, \ldots, z_n)) \\ &= (x_1, \ldots, x_n) + (y_1 + z_1, \ldots, y_n + z_n) \\ &= (x_1 + (y_1 + z_1), \ldots, x_n + (y_n + z_n)) \\ &= ((x_1 + y_1) + z_1, \ldots, (x_n + y_n) + z_n) \\ &= (x_1 + y_1, \ldots, x_n + y_n) + (z_1, \ldots, z_n) \\ &= (\vec{u} + \vec{v}) + \vec{w}. \end{aligned}$$

3. Ya que la suma de números reales es una operación conmutativa,

$$\begin{aligned} \vec{u} + \vec{v} &= (x_1, \ldots, x_n) + (y_1, \ldots, y_n) \\ &= (x_1 + y_1, \ldots, x_n + y_n) \\ &= (y_1 + x_1, \ldots, y_n + x_n) \\ &= (y_1, \ldots, y_n) + (x_1, \ldots, x_n) \\ &= \vec{v} + \vec{u}. \end{aligned}$$

4. Como el neutro aditivo de $\mathbb{R}$ es 0, se tiene

$$\begin{aligned} \vec{u} + \vec{0}_{\mathbb{R}^n} &= (x_1, \ldots, x_n) + (\underbrace{0, \ldots, 0}_{n}) \\ &= (x_1 + 0, \ldots, x_n + 0) \\ &= (x_1, \ldots, x_n) \\ &= \vec{u}. \end{aligned}$$

5. Si x es un número real, existe su inverso aditivo, $-x$, que satisface $x + (-x) = 0$; luego, si $-\vec{u} = (-x_1, \ldots, -x_n)$, se tiene

$$\begin{aligned}\vec{u}+(-\vec{u}) &= (x_1,\ldots,x_n)+(-x_1,\ldots,-x_n)\\ &= (x_1+(-x_1),\ldots,(x_n+(-x_n))\\ &= \underbrace{(0,\ldots,0)}_{n}\\ &= \vec{0}_{\mathbb{R}^n}.\end{aligned}$$

6. Ya que el producto de números reales es también un número real,

$$\lambda\vec{u} = \lambda(x_1,\ldots,x_n) = (\lambda x_1,\ldots,\lambda x_n)\in\mathbb{R}^n.$$

7. Puesto que el producto de números reales es asociativo,

$$\begin{aligned}\lambda(\beta\vec{u}) &= \lambda(\beta(x_1,\ldots,x_n))\\ &= \lambda(\beta x_1,\ldots,\beta x_n)\\ &= (\lambda(\beta x_1),\ldots,\lambda(\beta x_n))\\ &= ((\lambda\beta)x_1,\ldots,(\lambda\beta)x_n)\\ &= (\lambda\beta)(x_1,\ldots,x_n)\\ &= (\lambda\beta)\vec{u}.\end{aligned}$$

8. Dado que el producto se distribuye con respecto a la suma en los números reales,

$$\begin{aligned}(\lambda+\beta)\vec{u} &= ((\lambda+\beta)x_1,\ldots,(\lambda+\beta)x_n)\\ &= (\lambda x_1+\beta x_1,\ldots,\lambda x_n+\beta x_n)\\ &= (\lambda x_1,\ldots,\lambda x_n)+(\beta x_1,\ldots,\beta x_n)\\ &= \lambda(x_1,\ldots,x_n)+\beta(x_1,\ldots,x_n)\\ &= \lambda\vec{u}+\beta\vec{u}.\end{aligned}$$

9. Como el producto de números reales se distribuye con respecto a la suma, se tiene

$$\begin{aligned}\lambda(\vec{u}+\vec{v}) &= \lambda((x_1,\ldots,x_n)+(y_1,\ldots,y_n))\\ &= \lambda(x_1+y_1,\ldots,x_n+y_n)\\ &= (\lambda(x_1+y_1),\ldots,\lambda(x_n+y_n))\\ &= (\lambda x_1+\lambda y_1,\ldots,\lambda x_n+\lambda y_n)\\ &= (\lambda x_1,\ldots,\lambda x_n)+(\lambda y_1,\ldots,\lambda y_n)\\ &= \lambda(x_1,\ldots,x_n)+\beta(y_1,\ldots,y_n)\\ &= \lambda\vec{u}+\lambda\vec{v}.\end{aligned}$$

10. Puesto que $1x = x\ \forall x\in\mathbb{R}$,

$$\begin{aligned}1\vec{u} &= 1(x_1,\ldots,x_n)\\ &= (1x_1,\ldots,1x_n)\\ &= (x_1,\ldots,x_n)\\ &= \vec{u}. \quad ■\end{aligned}$$

15 Probar las propiedades 1 (simetría) y 2 (homogeneidad) del producto punto (cfr. pág. 123).

DEMOSTRACIÓN ■ Sea $\vec{u} = (x_1, \ldots, x_n)$, $\vec{v} = (y_1, \ldots, y_n)$ cualquier par de vectores en $\mathbb{R}^n$ y $\lambda \in \mathbb{R}$. Entonces:

1. (Simetría). Dado que el producto de números reales es conmutativo, se tiene

$$\begin{aligned}\vec{u} \cdot \vec{v} &= (x_1, \ldots, x_n) \cdot (y_1, \ldots, y_n)\\ &= \sum_{i=1}^{n} x_i y_i\\ &= \sum_{i=1}^{n} y_i x_i\\ &= (y_1, \ldots, y_n) \cdot (x_1, \ldots, x_n)\\ &= \vec{v} \cdot \vec{u}.\end{aligned}$$

2. (Homogeneidad). Puesto que el producto de números reales es asociativo y distributivo,

$$\begin{aligned}\vec{u} \cdot (\lambda \vec{v}) &= (x_1, \ldots, x_n) \cdot (\lambda(y_1, \ldots, y_n))\\ &= (x_1, \ldots, x_n) \cdot (\lambda y_1, \ldots, \lambda y_n)\\ &= \sum_{i=1}^{n} x_i(\lambda y_i)\\ &= \sum_{i=1}^{n} \lambda(x_i y_i)\\ &= \lambda \sum_{i=1}^{n} x_i y_i\\ &= \lambda(\vec{u} \cdot \vec{v}). \quad \blacksquare\end{aligned}$$

16 Hallar el ángulo entre $\vec{u} = (-1, 3, 4, -2)$ y $\vec{w} = (1, 0, 4, -2)$.

Solución

$$\begin{aligned}\theta &= \arccos\left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \, \|\vec{v}\|}\right)\\ &= \arccos\left(\frac{(-1,3,4,-2) \cdot (1,0,4,-2)}{\|(-1,3,4,-2)\| \, \|(1,0,4,-2)\|}\right)\\ &= \arccos\left(\frac{19}{\sqrt{30}\sqrt{21}}\right)\\ &\approx 0.71212\\ &\approx 40.801^\circ. \quad ✔\end{aligned}$$

17 Demostrar el recíproco del teorema de Pitágoras, es decir, si $\vec{u}, \vec{v} \in \mathbb{R}^n$ y $\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$, entonces $\vec{u}$ y $\vec{v}$ son ortogonales.

DEMOSTRACIÓN ■

$$\begin{aligned}\|\vec{u}\|^2 + \|\vec{v}\|^2 &= \|\vec{u} + \vec{v}\|^2\\ &= (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v})\\ &= (\vec{u} + \vec{v}) \cdot \vec{u} + (\vec{u} + \vec{v}) \cdot \vec{v}\\ &= \vec{u} \cdot \vec{u} + \vec{v} \cdot \vec{u} + \vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{v}\\ &= \|\vec{u}\|^2 + 2(\vec{u} \cdot \vec{v}) + \|\vec{v}\|^2 \Rightarrow\\ 2(\vec{u} \cdot \vec{v}) &= 0 \Rightarrow\\ \vec{u} \cdot \vec{v} &= 0. \quad \blacksquare\end{aligned}$$

18 Sean $\vec{u}_i$, $i = 1,\ldots,k$, vectores de $\mathbb{R}^n$ que son ortogonales entre sí; es decir, $\vec{u}_i \cdot \vec{u}_j = 0$ si $i \neq j$.

(a) Demostrar por inducción la generalización del teorema de Pitágoras; esto es,

$$\|\vec{u}_1 + \vec{u}_2 + \cdots + \vec{u}_k\|^2 = \|\vec{u}_1\|^2 + \|\vec{u}_2\|^2 + \cdots + \|\vec{u}_k\|^2.$$

(b) Si además los vectores $\vec{u}_i$ no son nulos, probar, utilizando el primer inciso, que los vectores $\vec{u}_i$ son L.I.

DEMOSTRACIÓN ■ *(a)* El resultado claramente es válido para $k = 1$. Suponga que la afirmación es cierta para $k-1$ vectores. Puesto que los vectores son ortogonales entre sí, se tiene

$$\left(\sum_{i=1}^{k-1} \vec{u}_i\right) \cdot \vec{u}_k = \sum_{i=1}^{k-1} (\vec{u}_i \cdot \vec{u}_k) = 0;$$

luego los vectores $\sum_{i=1}^{k-1} \vec{u}_i$ y $\vec{u}_k$ son ortogonales, así que, por el teorema de Pitágoras (teorema 3.2, pág. 126),

$$\|(\vec{u}_1 + \cdots + \vec{u}_{k-1}) + \vec{u}_k\|^2 = \|\vec{u}_1 + \cdots + \vec{u}_{k-1}\|^2 + \|\vec{u}_k\|^2$$

y por hipótesis de inducción

$$\|\vec{u}_1 + \cdots + \vec{u}_{k-1}\|^2 = \|\vec{u}_1\|^2 + \cdots + \|\vec{u}_{k-1}\|^2;$$

por lo que

$$\begin{aligned} \|\vec{u}_1 + \cdots + \vec{u}_{k-1} + \vec{u}_k\|^2 &= \|\vec{u}_1 + \cdots + \vec{u}_{k-1}\|^2 + \|\vec{u}_k\|^2 \\ &= \|\vec{u}_1\|^2 + \cdots + \|\vec{u}_{k-1}\|^2 + \|\vec{u}_k\|^2. \end{aligned}$$

(b) Sean α_i, $i = 1,\ldots,k$, escalares tales que

$$\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_k\vec{u}_k = \vec{0}_{\mathbb{R}^n}.$$

Puesto que $(\alpha_i\vec{u}_i) \cdot (\alpha_j\vec{u}_j) = \alpha_i\alpha_j(\vec{u}_i \cdot \vec{u}_j) = \alpha_i\alpha_j \cdot 0 = 0$ si $i \neq j$, se sigue que los vectores $\alpha_i\vec{u}_i$ son ortogonales entre sí. Por el inciso precedente, se tiene

$$\begin{aligned} 0 &= \|\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_k\vec{u}_k\|^2 \\ &= \alpha_1^2\|\vec{u}_1\|^2 + \alpha_2^2\|\vec{u}_2\|^2 + \cdots + \alpha_k^2\|\vec{u}_k\|^2; \end{aligned}$$

por tanto,

$$\alpha_i^2\|\vec{u}_i\|^2 = 0 \quad \text{para todo } i.$$

Ya que los vectores $\vec{u}_i$ no son nulos, se deduce que $\alpha_i = 0$ para todo $i = 1,\ldots,k$, y por ende los vectores $\vec{u}_i$ son L.I. ■

19 Probar la propiedad 4 (desigualdad triangular) del valor absoluto de la página 127:

$$|x+y| \leq |x| + |y| \quad \forall x,y \in \mathbb{R}.$$

DEMOSTRACIÓN ■ Sean $x, y \in \mathbb{R}$. Dado que $a \leq |a|$ para todo número real a, se tiene

$$\begin{aligned}(x+y)^2 &= x^2+2xy+y^2\\ &\leq |x|^2+2|x||y|+|y|^2\\ &= (|x|+|y|)^2,\end{aligned}$$

de donde

$$\sqrt{(x+y)^2} \leq \sqrt{(|x|+|y|)^2};$$

y puesto que $|a| = \sqrt{a^2}$, la anterior desigualdad implica

$$|x+y| \leq |x|+|y|. \quad ■$$

20 Encontrar la ecuación de un plano que sea perpendicular al plano $3x-2y+4z=12$ y pase por el punto $(-1,2,1)$.

Solución Si $y=0$ y $z=0$ en la ecuación del plano dado, entonces $x=4$; así que el punto $(4,0,0)$ pertenece a este plano. Si $x=0$ y $y=0$ en la ecuación del plano dado, entonces $z=3$; por tanto, $(0,0,3)$ pertenece a este plano. Así que el vector $(4,0,-3)$ es un vector perpendicular al plano que se está buscando. Entonces

$$4(x+1)+0(y-2)-3(z-1)=0$$

o, equivalentemente,

$$4x-3z=-7$$

resuelve este problema. ✔

Espacios vectoriales

21 Probar que $\mathfrak{M}_{m\times n}$ con las operaciones usuales de suma de matrices y multiplicación de un escalar por una matriz es un espacio vectorial.

DEMOSTRACIÓN ■ Sean $A=[a_{ij}]$, $B=[b_{ij}]$ y $C=[c_{ij}]$ matrices arbitrarias de tamaño $m\times n$; y $\alpha,\beta\in\mathbb{R}$ cualquier par de números. Entonces:

1. Puesto que la suma de números reales da como resultado un número real, se tiene $a_{ij}+b_{ij}\in\mathbb{R}$ para todo $i=1,\ldots,m$ y para todo $j=1,\ldots,n$; por tanto,

$$A+B=[a_{ij}+b_{ij}]\in\mathfrak{M}_{m\times n}.$$

2. Dado que la suma de números reales es asociativa, $(a_{ij}+b_{ij})+c_{ij}=a_{ij}+(b_{ij}+c_{ij})$ para todo $i=1,\ldots,m$ y para todo $j=1,\ldots,n$; por tanto,

$$(A+B)+C=A+(B+C).$$

3. Ya que la suma de números reales es conmutativa, $a_{ij}+b_{ij}=b_{ij}+a_{ij}$ para todo $i=1,\ldots,m$ y para todo $j=1,\ldots,n$; por ende,

$$A+B=B+A.$$

4. Como $x+0=x \quad \forall x\in\mathbb{R}$, $a_{ij}+0=a_{ij}$, para todo $i=1,\ldots,m$ y para todo $j=1,\ldots,n$; luego

$$A+\mathscr{O}=A.$$

5. Puesto que para cada $x\in\mathbb{R}$ existe $-x\in\mathbb{R}$ tal que $x+(-x)=0$, se tiene que si $-A=[-a_{ij}]$, entonces

$$A+(-A)=\mathscr{O}.$$

6. La multiplicación de números reales da como resultado un número real; por tanto, $\alpha a_{ij}\in\mathbb{R}$ para todo $i=1,\ldots,m$ y para todo $j=1,\ldots,n$; luego

$$\alpha A=[\alpha a_{ij}]\in\mathfrak{M}_{m\times n}.$$

7. La multiplicación de números reales es asociativa; por tanto,

$$\begin{aligned}(\alpha\beta)A &= [(\alpha\beta)a_{ij}]\\ &= [\alpha(\beta a_{ij})]\\ &= \alpha[\beta a_{ij}]\\ &= \alpha(\beta A).\end{aligned}$$

8. La multiplicación de números reales se distribuye con respecto a la suma, entonces,

$$\begin{aligned}\alpha(A+B) &= \alpha([a_{ij}]+[b_{ij}])\\ &= \alpha[a_{ij}+b_{ij}]\\ &= [\alpha(a_{ij}+b_{ij})]\\ &= [\alpha a_{ij}+\alpha b_{ij}]\\ &= [\alpha a_{ij}]+[\alpha b_{ij}]\\ &= \alpha[a_{ij}]+\alpha[b_{ij}]\\ &= \alpha A+\alpha B.\end{aligned}$$

9. Puesto que la multiplicación de números reales se distribuye con respecto a la suma,

$$\begin{aligned}(\alpha+\beta)A &= (\alpha+\beta)[a_{ij}]\\ &= [(\alpha+\beta)a_{ij}]\\ &= [\alpha a_{ij}+\beta a_{ij}]\\ &= [\alpha a_{ij}]+[\beta a_{ij}]\\ &= \alpha[a_{ij}]+\beta[a_{ij}]\\ &= \alpha A+\beta A.\end{aligned}$$

10. Ya que $1x = x$ para todo $x \in \mathbb{R}$, se tiene

$$\begin{aligned} 1A &= 1[a_{ij}] \\ &= [1a_{ij}] \\ &= [a_{ij}] \\ &= A. \quad \blacksquare \end{aligned}$$

22 Sea $\mathbf{E}$ el conjunto de matrices 3×3 con la operación usual de multiplicación de un escalar por una matriz pero con la suma definida en la forma $A \oplus B = \mathscr{O}$, donde $\mathscr{O}$ es la matriz cero 3×3. Determinar si $\mathbf{E}$ es un espacio vectorial con estas operaciones. En el caso de que $\mathbf{E}$ no sea un espacio vectorial indicar los axiomas que no cumplen.

Solución Sean $A, B, C \in \mathbf{E}$ y $\alpha, \beta \in \mathbb{R}$. (1) Claramente la suma $\oplus$ es cerrada. (2) $A \oplus (B \oplus C) = A \oplus \mathscr{O} = \mathscr{O} = \mathscr{O} + C = (A \oplus B) \oplus C$. (3) $A \oplus B = \mathscr{O} = B \oplus A$. (4) Si A es la matriz 3×3 con $a_{11} = 1$ y las demás componentes iguales a cero, entonces $A \oplus B = \mathscr{O} \neq A \quad \forall B \in \mathbf{E}$, luego $\mathbf{E}$ no tiene neutro aditivo. (5) Por el inciso anterior, no tiene sentido el concepto de inverso aditivo. (6), (7) y (10) se cumplen porque es la multiplicación usual de un escalar por una matriz. (8) $\alpha(A \oplus B) = \alpha\mathscr{O} = \alpha A \oplus \alpha B$. (9) Sea A como en (4), entonces $(1+2)A$ es la matriz 3×3 con $a_{11} = 3$ y las demás componentes nulas; mientras que $1A \oplus 2A = \mathscr{O} \neq (1+2)A$. En resumen, los axiomas 4, 5 y 9 no se cumplen pero los demás sí. $\mathbf{E}$ no es un espacio vectorial con estas operaciones. ✔

23 Sean $\mathbf{E}$ un espacio vectorial y $S \subset \mathbf{E}$. Mostrar que $S < \mathbf{E}$ (S es un subespacio de $\mathbf{E}$) si y sólo si se cumplen las dos siguientes condiciones:

(i) $S \neq \emptyset$.

(ii) $\alpha\vec{u} + \beta\vec{v} \in S \quad \forall \vec{u}, \vec{v} \in S, \forall \alpha, \beta \in \mathbb{R}$.

DEMOSTRACIÓN ■ ($\Rightarrow$) Supongamos que $S < \mathbf{E}$. Entonces $S \neq \emptyset$ pues $\vec{0}_{\mathbf{E}} \in S$. Sean $\vec{u}, \vec{v} \in S$ y $\alpha, \beta \in \mathbb{R}$; ya que $S < \mathbf{E}$, la multiplicación de vectores por escalares y la suma de vectores es cerrada en S, luego $\alpha\vec{u} \in S$ y $\beta\vec{v} \in S$; por tanto $\alpha\vec{u} + \beta\vec{v} \in S$.

($\Leftarrow$) Supongamos que se cumplen *(i)* y *(ii)*; entonces, por *(i)* existe $\vec{u} \in S$; y por *(ii)* $\vec{0}_{\mathbf{E}} = 0\vec{u} \in S$. Si $\vec{w}, \vec{v} \in S$, y $\alpha \in \mathbb{R}$, por *(ii)* $\vec{w} + \vec{v} \in S$ y $\alpha\vec{w} \in S$. Por tanto $S < \mathbf{E}$. ■

24 Sea S el conjunto de todas las funciones, y, derivables hasta el orden dos en todos los puntos de $\mathbb{R}$ que satisfacen la relación

$$y''(x) - y'(x) - 6y(x) = 0 \quad \forall x \in \mathbb{R}.$$

Mostrar que S es un espacio vectorial.

DEMOSTRACIÓN ■ Puesto que $S \subset \mathscr{F}(\mathbb{R})$, basta probar que $S < \mathscr{F}(\mathbb{R})$.

(a) Ya que θ, la función constante cero, satisface $\theta'(x) = 0$ y $\theta''(x) = 0$ para todo $x \in \mathbb{R}$, se tiene $\theta \in S$.

(b) Si $f, g \in S$, entonces $f''(x) - f'(x) - 6f(x) = 0$ y $g''(x) - g'(x) - 6g(x) = 0$ para todo $x \in \mathbb{R}$. Luego, para cada $x \in \mathbb{R}$,

$$\begin{aligned}(f+g)''(x)-(f+g)'(x)-6(f+g)(x) &= f''(x)-f'(x)-6f(x)\\ &\quad +g''(x)-g'(x)-6g'(x)=0;\end{aligned}$$

por tanto $f+g\in S$.

(c) Si $\alpha\in\mathbb{R}$ y $f\in S$, entonces $f''(x)-f'(x)-6f(x)=0$ para todo $x\in\mathbb{R}$, así que

$$\begin{aligned}(\alpha f)''(x)-(\alpha f)'(x)-6(\alpha f)'(x) &= \alpha(f''(x)-f'(x)-6f(x))\\ &= \alpha 0\\ &= 0 \quad \forall x\in\mathbb{R}.\end{aligned}$$

Por tanto, $\alpha f\in S$.

De 1, 2 y 3, $S<\mathscr{F}(\mathbb{R})$ y por ende es un espacio vectorial. ■

25 Una sucesión (a_n) de números reales es acotada si existe $M>0$ tal que $|a_n|\le M$ para todo n, se dice entonces que M es una cota para la sucesión. Sea S el subconjunto, del espacio de sucesiones $\mathbb{R}^\infty$, formado por todas las sucesiones que son acotadas. Probar que $S<\mathbb{R}^\infty$.

DEMOSTRACIÓN ■ Sean $(a_n),(b_n)\in S$, $\alpha\in\mathbb{R}$.

1. Claramente la sucesión constante cero pertenece a S (cualquier número positivo es una cota para esta sucesión).

2. Sean $M_1,M_2>0$ cotas de las sucesiones (a_n) y (b_n), respectivamente; esto es, $|a_n|\le M_1$ y $|b_n|\le M_2$ para todo n. Entonces, por la desigualdad triangular para el valor absoluto de números reales,

$$|a_n+b_n|\le|a_n|+|b_n|\le M_1+M_2$$

para todo n, entonces $M_1+M_2>0$ es una cota para la sucesión (a_n+b_n); lo cual implica $(a_n)+(b_n)\in S$.

3. Para cada n,

$$|\alpha a_n|=|\alpha||a_n|\le(|\alpha|+1)M_1.$$

Así, $(|\alpha|+1)M_1>0$ es una cota para la sucesión (αa_n); por tanto, $\alpha(a_n)\in S$.

De 1, 2 y 3, $S<\mathbb{R}^\infty$. ■

26 (**Suma de subespacios**). Sean $\mathbf{E}$ un espacio vectorial y R,S un par de subespacios. Se define el conjunto, suma de los subespacios R y S, como

$$R+S=\{\vec{u}+\vec{v}\,|\,\vec{u}\in R,\vec{v}\in S\}.$$

Mostrar que $R+S<\mathbf{E}$.

DEMOSTRACIÓN ■ (1) Dado que R y S son subespacios de $\mathbf{E}$, $\vec{0}_{\mathbf{E}} \in R \cap S$; por lo que, $\vec{0}_{\mathbf{E}} = \vec{0}_{\mathbf{E}} + \vec{0}_{\mathbf{E}} \in R + S$. (2) Si $\vec{u}_1 + \vec{v}_1, \vec{u}_2 + \vec{v}_2 \in R + S$, entonces, $\vec{u}_1, \vec{u}_2 \in R$ y $\vec{v}_1, \vec{v}_2 \in S$; por tanto, $\vec{u}_1 + \vec{u}_2 \in R$ y $\vec{v}_1 + \vec{v}_2 \in S$; así que

$$(\vec{u}_1 + \vec{v}_1) + (\vec{u}_2 + \vec{v}_2) = (\vec{u}_1 + \vec{u}_2) + (\vec{v}_1 + \vec{v}_2) \in R + S.$$

(3) Si $\alpha \in \mathbb{R}$ y $\vec{u} + \vec{v} \in R + S$, entonces $\vec{u} \in R$ y $\vec{v} \in S$, por lo que $\alpha\vec{u} \in R$ y $\alpha\vec{v} \in S$; luego

$$\alpha(\vec{u} + \vec{v}) = \alpha\vec{u} + \alpha\vec{v} \in R + S.$$

De 1, 2 y 3, $R + S < \mathbf{E}$. ■

27 (**Suma directa**). Dado un espacio vectorial $\mathbf{E}$ y k subespacios $S_i \subset \mathbf{E}$, se dice que $\mathbf{E}$ es la suma directa de estos subespacios y se escribe

$$\mathbf{E} = S_1 \oplus S_2 \oplus \cdots \oplus S_k$$

si:

- $\mathbf{E} = S_1 + S_2 + \cdots + S_k$; esto es, todo elemento $\vec{w} \in \mathbf{E}$ tiene la forma

$$\vec{w} = \vec{u}_1 + \vec{u}_2 + \cdots + \vec{u}_k,$$

con $\vec{u}_i \in S_i$ para cada $i = 1, 2, \ldots, k$.
- $S_i \cap (S_1 + S_2 + \cdots + S_{i-1}) = \{\vec{0}_{\mathbf{E}}\} \quad \forall i,\ 2 \le i \le k$.

Probar que $\mathbf{E}$ es la suma directa de los subespacios $S_1, \ldots, S_k$ si y sólo si todo vector $\vec{w} \in \mathbf{E}$ se puede escribir de manera única como $\vec{w} = \vec{u}_1 + \cdots + \vec{u}_k$, con $\vec{u}_i \in S_i, i = 1, \ldots, k$.

DEMOSTRACIÓN ■ ($\Rightarrow$) Si $\mathbf{E}$ es la suma directa de los subespacios S_i, y $\vec{w} = \sum_{i=1}^{k} \vec{u}_i = \sum_{i=1}^{k} \vec{v}_i$ con $\vec{u}_i, \vec{v}_i \in S_i$, entonces

$$\vec{u}_k - \vec{v}_k = -\sum_{i=1}^{k-1} (\vec{u}_i - \vec{v}_i) \in S_k \cap (S_1 + \cdots + S_{k-1}) = \{\vec{0}_{\mathbf{E}}\}$$

por tanto $\vec{u}_k = \vec{v}_k$. Luego,

$$\vec{u}_{k-1} - \vec{v}_{k-1} = -\sum_{i=1}^{k-2} (\vec{u}_i - \vec{v}_i) \in S_{k-1} \cap (S_1 + \cdots + S_{k-2}) = \{\vec{0}_{\mathbf{E}}\},$$

de donde $\vec{u}_{k-1} = \vec{v}_{k-1}$. De manera análoga se prueba $\vec{u}_i = \vec{v}_i$ para los otros índices.
($\Leftarrow$) Claramente $\mathbf{E} = S_1 + \cdots + S_k$. Sea $\vec{u} \in (S_1 + \cdots + S_{j-1}) \cap S_j$, entonces existen $\vec{u}_i \in S_i$ tales que $\vec{u} = \vec{u}_1 + \cdots + \vec{u}_{j-1}$. Luego

$$\vec{0}_{\mathbf{E}} = \vec{u}_i + \cdots + \vec{u}_{j-1} + (-\vec{u}) + \underbrace{\vec{0}_{\mathbf{E}} + \cdots + \vec{0}_{\mathbf{E}}}_{k-j};$$

y por la unicidad de los términos, se desprende que $\vec{u}_i = \vec{0}_{\mathbf{E}} = \vec{u}\ \forall i$.

28 Sea $\mathfrak{M}_n$ el espacio de matrices cuadradas de orden n; del ejemplo 3.25 (pág. 143) el conjunto S_1, de matrices cuadradas simétricas de orden n, es un subespacio de $\mathfrak{M}_n$.

(a) Sea S_2 el conjunto de matrices cuadradas A de orden n que son antisimétricas; esto es, $A^t = -A$. Probar que S_2 es un subespacio de $\mathfrak{M}_n$.

(b) Demostrar que $\mathfrak{M}_n = S_1 \oplus S_2$.

DEMOSTRACIÓN ■ *(a)* Claramente $\mathscr{O} \in S_2$. Si $A, B \in S_2$ y $\alpha, \beta \in \mathbb{R}$, entonces

$$\begin{aligned}(\alpha A + \beta B)^t &= \alpha A^t + \beta B^t \\ &= \alpha(-A) + \beta(-B) \\ &= -(\alpha A + \beta B).\end{aligned}$$

Lo cual prueba que $S_2 < \mathfrak{M}_n$.

(b) Sean $A \in \mathfrak{M}_n$, $M_1 = \frac{1}{2}A + \frac{1}{2}A^t$ y $M_2 = \frac{1}{2}A - \frac{1}{2}A^t$. Claramente $A = M_1 + M_2$, y puesto que

$$\begin{aligned}M_1^t &= (\tfrac{1}{2}A + \tfrac{1}{2}A^t)^t = \tfrac{1}{2}A^t + \tfrac{1}{2}A = M_1 \text{ y} \\ M_2^t &= (\tfrac{1}{2}A - \tfrac{1}{2}A^t)^t = \tfrac{1}{2}A^t - \tfrac{1}{2}A = -M_2,\end{aligned}$$

se tiene que $M_1 \in S_1$ y $M_2 \in S_2$; con lo que se ha probado que $\mathfrak{M}_n = S_1 + S_2$. Si $M \in S_1 \cap S_2$, entonces $M^t = M$ y $M^t = -M$, de donde $M = -M$; por tanto, $M = \mathscr{O}$; es decir, $S_1 \cap S_2 = \{\mathscr{O}\}$. Luego $\mathfrak{M}_n = S_1 \oplus S_2$. ■

29 (**Espacio cociente**). Sean $\mathbf{E}$ un espacio vectorial y $S \subset \mathbf{E}$ un subespacio.

(a) Se define la siguiente relación entre los elementos de $\mathbf{E}$:

$$\vec{u} \sim \vec{v} \Leftrightarrow \vec{u} - \vec{v} \in S.$$

Probar que esta relación es de equivalencia; esto es, *(i)* $\vec{u} \sim \vec{u} \quad \forall \vec{u} \in \mathbf{E}$ (reflexividad); *(ii)* si $\vec{u}, \vec{v} \in \mathbf{E}$, $\vec{u} \sim \vec{v} \Rightarrow \vec{v} \sim \vec{u}$ (simetría); *(iii)* si $\vec{u}, \vec{v}, \vec{w} \in \mathbf{E}$, $\vec{u} \sim \vec{v}, \vec{v} \sim \vec{w} \Rightarrow \vec{u} \sim \vec{w}$ (transitividad).

(b) Para cada $\vec{u} \in \mathbf{E}$, el símbolo $[\vec{u}]$ representa la clase de equivalencia del vector $\vec{u}$; es decir, el conjunto de vectores que están relacionados con $\vec{u}$. Probar que dos clases de equivalencia o son disjuntas (su intersección es vacía) o son iguales.

(c) Si

$$\mathbf{E}/S = \{[\vec{u}] \,|\, \vec{u} \in \mathbf{E}\},$$

se definen

$$[\vec{u}] + [\vec{v}] = [\vec{u} + \vec{v}]$$

y

$$\alpha[\vec{u}] = [\alpha\vec{u}]$$

para cada $[\vec{u}], [\vec{v}] \in \mathbf{E}/S$ y para cada $\alpha \in \mathbb{R}$. Probar que las operaciones $[\vec{u}] + [\vec{v}]$ y $\alpha[\vec{u}]$ están bien definidas; esto es, que no dependen de los representantes de cada clase de equivalencia; o sea, que si $[\vec{u}] = [\vec{u}']$ y $[\vec{v}] = [\vec{v}']$, entonces $[\vec{u} + \vec{v}] = [\vec{u}' + \vec{v}']$ y $[\alpha\vec{u}] = [\alpha\vec{u}']$.

(d) Probar que $\mathbf{E}/S$, junto con las operaciones definidas en el inciso precedente, es un espacio vectorial (llamado espacio cociente del subespacio S).

DEMOSTRACIÓN ■ *(a)* Sean $\vec{u},\vec{v},\vec{w}$ vectores arbitrarios de $\mathbf{E}$.

(i) $\vec{u}-\vec{u}=\vec{0}_{\mathbf{E}}\in S$, porque $S<\mathbf{E}$. Por tanto $\vec{u}\sim\vec{u}$.

(ii) $\vec{u}\sim\vec{v}\Rightarrow\vec{u}-\vec{v}\in S \overset{(\text{porque } S<\mathbf{E})}{\Rightarrow} \vec{v}-\vec{u}=-(\vec{u}-\vec{v})\in S\Rightarrow\vec{v}\sim\vec{u}$.

(iii) $\vec{u}\sim\vec{v}$ y $\vec{v}\sim\vec{w}\Rightarrow\vec{u}-\vec{v},\vec{v}-\vec{w}\in S \overset{(\text{porque } S<\mathbf{E})}{\Rightarrow} \vec{u}-\vec{w}=(\vec{u}-\vec{v})+(\vec{v}-\vec{w})\in S\Rightarrow\vec{u}\sim\vec{w}$.

(b) Si $\vec{u},\vec{v}\in\mathbf{E}$ y $[\vec{u}]\cap[\vec{v}]\neq\emptyset$, entonces existe $\vec{w}\in\mathbf{E}$ tal que $\vec{u}\sim\vec{w}$ y $\vec{w}\sim\vec{v}$; lo cual implica $\vec{u}\sim\vec{v}$; por tanto, $[\vec{u}]=[\vec{v}]$.

(c) Sean $\vec{u},\vec{u}',\vec{v},\vec{v}'\in\mathbf{E}$ tales que $\vec{u}\sim\vec{u}'$ y $\vec{v}\sim\vec{v}'$ (esto es, $[\vec{u}]=[\vec{u}']$, $[\vec{v}]=[\vec{v}']$). Entonces $\vec{u}-\vec{u}'\in S$ y $\vec{v}-\vec{v}'\in S$, y puesto que $S<\mathbf{E}$, se tiene $(\vec{u}+\vec{v})-(\vec{u}'+\vec{v}')\in S$, lo cual implica $\vec{u}+\vec{v}\sim\vec{u}'+\vec{v}'$; por tanto, $[\vec{u}+\vec{v}]=[\vec{u}'+\vec{v}']$. Si $\vec{u}\sim\vec{u}'$, entonces $\vec{u}-\vec{u}'\in S$ y, ya que este último es subespacio, se tiene $\alpha\vec{u}-\alpha\vec{u}'\in S$; por lo que $[\alpha\vec{u}]=[\alpha\vec{u}']$.

(d) Sean $\vec{u},\vec{v}$ y $\vec{w}$ vectores arbitrarios de $\mathbf{E}$ y α,β cualquier par de números reales.

1. Claramente $[\vec{u}]+[\vec{v}]=[\vec{u}+\vec{v}]\in S$.

2. $$\begin{aligned}[\vec{u}]+([\vec{v}]+[\vec{w}]) &= [\vec{u}]+[\vec{v}+\vec{w}]\\ &= [\vec{u}+(\vec{v}+\vec{w})]\\ &= [(\vec{u}+\vec{v})+\vec{w}]\\ &= [\vec{u}+\vec{v}]+[\vec{w}].\end{aligned}$$

3. $$\begin{aligned}[\vec{u}]+[\vec{v}] &= [\vec{u}+\vec{v}]\\ &= [\vec{v}+\vec{u}]\\ &= [\vec{v}]+[\vec{u}].\end{aligned}$$

4. (Note que $[\vec{0}_{\mathbf{E}}]=S$.) $$\begin{aligned}[\vec{u}]+[\vec{0}_{\mathbf{E}}] &= [\vec{u}+\vec{0}_{\mathbf{E}}]\\ &= [\vec{u}].\end{aligned}$$

5. $$[\vec{u}]+[-\vec{u}]=[\vec{0}_{\mathbf{E}}].$$

6. Por definición, $\alpha[\vec{u}]\in\mathbf{E}/S$.

7. $$\begin{aligned}\alpha(\beta[\vec{u}]) &= \alpha[\beta\vec{u}]\\ &= [\alpha(\beta\vec{u})]\\ &= [(\alpha\beta)\vec{u}]\\ &= (\alpha\beta)[\vec{u}].\end{aligned}$$

8. $$\begin{aligned}(\alpha+\beta)[\vec{u}] &= [(\alpha+\beta)\vec{u}]\\ &= [\alpha\vec{u}+\beta\vec{u}]\\ &= [\alpha\vec{u}]+[\beta\vec{u}]\\ &= \alpha[\vec{u}]+\beta[\vec{u}].\end{aligned}$$

9.
$$\begin{aligned}\alpha([\vec{u}]+[\vec{v}]) &= \alpha[\vec{u}+\vec{v}]\\ &= [\alpha(\vec{u}+\vec{v})]\\ &= [\alpha\vec{u}+\alpha\vec{v}]\\ &= [\alpha\vec{u}]+[\alpha\vec{v}]\\ &= \alpha[\vec{u}]+\alpha[\vec{v}].\end{aligned}$$

10.
$$\begin{aligned}1[\vec{u}] &= [1\vec{u}]\\ &= [\vec{u}].\end{aligned}$$

Los incisos 1 a 10 implican que $\mathbf{E}/S$ es un espacio vectorial. ■

30 Sea Λ un conjunto no vacío y S_λ, $\lambda\in\Lambda$, una familia de subespacios de un espacio vectorial $\mathbf{E}$. Sea

$$S=\bigcap_{\lambda\in\Lambda}S_\lambda=\{\vec{u}\in\mathbf{E}\,|\,\vec{u}\in S_\lambda \text{ para todo } \lambda\in\Lambda\}.$$

Probar que S es un subespacio de $\mathbf{E}$.

DEMOSTRACIÓN ■ *(a)* Puesto que $S_\lambda<\mathbf{E}$ $\forall\lambda\in\Lambda$, $\vec{0}_{\mathbf{E}}\in S_\lambda$ para todo λ; por tanto $\vec{0}_{\mathbf{E}}\in S$.

(b) Sean $\vec{u},\vec{v}\in S$ y $\alpha\in\mathbb{R}$.

(i) Ya que $\vec{u},\vec{v}\in S_\lambda<\mathbf{E}$ para todo λ, se tiene $\vec{u}+\vec{v}\in S_\lambda<\mathbf{E}$ para todo λ, y en consecuencia $\vec{u}+\vec{v}\in S$.

(ii) $\vec{u}\in S\Rightarrow\vec{u}\in S_\lambda$ para todo λ y como $S_\lambda<\mathbf{E}$, $\alpha\vec{u}\in S_\lambda\,\forall\lambda\in\Lambda$; por tanto, $\alpha\vec{u}\in S$.

De *(a)* y *(b)*, $S<\mathbf{E}$. ■

31 (**Subespacio generado por un conjunto**). Sean $\mathbf{E}$ un espacio vectorial y $C\subset\mathbf{E}$. Se denota por $\mathscr{L}(C)$ la intersección de todos los subespacios de $\mathbf{E}$ que contienen a C. Por el inciso anterior $\mathscr{L}(C)$ es un subespacio de $\mathbf{E}$, llamado *el subespacio generado por el conjunto* C. Probar que $\mathscr{L}(C)$ es el menor subespacio que contiene a C; es decir, que cualquier subespacio que contenga a C contiene también a $\mathscr{L}(C)$.

DEMOSTRACIÓN ■ Sea $W<\mathbf{E}$ que contiene a C, entonces

$$W\in\{V\,|\,V \text{ es subespacio de } \mathbf{E} \text{ y } C\subset V\}=\Lambda;$$

luego, si $\vec{u}\in\mathscr{L}(C)$, que es la intersección de todos los elementos de Λ, $\vec{u}\in W$ porque $W\in\Lambda$; por tanto, $\mathscr{L}(C)\subset W$. ■

32 Sea $\mathscr{L}(C)$ como en el inciso anterior.

(a) Determinar $\mathscr{L}(\emptyset)$.

(b) Probar que si $C\neq\emptyset$, entonces $\mathscr{L}(C)$ es el conjunto de todas las posibles combinaciones lineales de elementos de C.

(c) Mostrar que si $C=\{\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_n\}$, con los $\vec{u}_i\in\mathbf{E}$, entonces

$$\mathscr{L}(C)=\text{gn}(\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_n).$$

Solución *(a)* El subespacio trivial $\{\vec{0}_{\mathbf{E}}\}$ está contenido en todo subespacio de $\mathbf{E}$, y $\emptyset \subset \{\vec{0}_{\mathbf{E}}\}$ (porque el conjunto vacío está contenido en cualquier conjunto); por tanto, $\mathscr{L}(\emptyset) = \{\vec{0}_{\mathbf{E}}\}$.

(b) Sea S el conjunto de todas las posibles combinaciones lineales de vectores en C; esto es, $\vec{v} \in S$ si y sólo si existen $\vec{u}_1, \ldots, \vec{u}_k \in C$ y $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, k$, tales que $\vec{v} = \sum_{i=1}^{k} \alpha_i \vec{u}_i$. Como C es no vacío, tiene por lo menos un elemento $\vec{u}$; por tanto, $\vec{0}_{\mathbf{E}} = 0\vec{u} \in S$. Si $\vec{v}_1 = \sum_{i=1}^{k} \alpha_i \vec{u}_i$ y $\vec{v}_2 = \sum_{i=1}^{m} \beta_i \vec{w}_i$ pertenecen a S, y $a, b \in \mathbb{R}$, entonces los $\vec{u}_i$ y los $\vec{w}_i$ pertenecen a C; por tanto,

$$\begin{aligned} a\vec{v}_1 + b\vec{v}_2 &= a\sum_{i=1}^{k} \alpha_i \vec{u}_i + b\sum_{i=1}^{m} \beta_i \vec{w}_i \\ &= \sum_{i=1}^{k} (a\alpha_i)\vec{u}_i + \sum_{i=1}^{m} (b\beta_i)\vec{w}_i \in S. \end{aligned}$$

Lo cual demuestra que $S < \mathbf{E}$. Claramente $C \subset S$ (si $\vec{u} \in C$, $\vec{u} = 1\vec{u} \in S$). Sea W un subespacio que contiene a C, entonces si $\vec{v} = \sum_{i=1}^{k} \alpha_i \vec{u}_i \in S$, los $\vec{u}_i \in C \subset W$; por tanto (ya que $W < \mathbf{E}$) $\vec{v} = \sum_{i=1}^{k} \alpha_i \vec{u}_i \in W$; luego $S \subset V$ para todo subespacio V que contiene a C, lo cual implica $S = \mathscr{L}(C)$.

(c) Es inmediata del inciso anterior y de la definición de $\text{gn}(\vec{u}_1, \ldots, \vec{u}_n)$. ✔

33 Encontrar condiciones sobre a, b, c para que el vector (a, b, c) pertenezca al espacio generado por

$$\vec{u}_1 = (1, -2, 1), \quad \vec{u}_2 = (2, -1, 3) \quad \text{y} \quad \vec{u}_3 = (-3, 0, -5).$$

¿$\mathbb{R}^3$ está generado por $\vec{u}_1$, $\vec{u}_2$, $\vec{u}_3$?

Solución $(a, b, c) \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ si el sistema $x_1\vec{u}_1 + x_2\vec{u}_2 + x_3\vec{u}_3 = (a, b, c)$ es consistente y ya que

$$\begin{aligned} \begin{bmatrix} 1 & 2 & -3 & a \\ -2 & -1 & 0 & b \\ 1 & 3 & -5 & c \end{bmatrix} &\sim \begin{bmatrix} 1 & 2 & -3 & a \\ 0 & 3 & -6 & 2a+b \\ 0 & 1 & -2 & c-a \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 2 & -3 & a \\ 0 & 1 & -2 & c-a \\ 0 & 3 & -6 & 2a+b \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 2 & -3 & a \\ 0 & 1 & -2 & c-a \\ 0 & 0 & 0 & 5a+b-3c \end{bmatrix}, \end{aligned}$$

esto sucede si $5a + b - 3c = 0$. Por tanto, $\mathbb{R}^3$ no está generado por $\vec{u}_1$, $\vec{u}_2$, $\vec{u}_3$; pues, por ejemplo, $(1, 1, 1) \notin \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ ya que $5(1) + (1) - 3(1) = 3 \neq 0$. ✔

34 Probar que $\text{gn}(1, x) = \text{gn}(2 + 3x, x)$.

DEMOSTRACIÓN ■

$$\begin{aligned} 1 &= a(2+3x) + bx \Rightarrow \\ 1 &= 2a + (3a+b)x \Rightarrow \\ a &= \frac{1}{2} \text{ y } b = -\frac{3}{2}; \end{aligned}$$

por tanto, $1 \in \text{gn}(2 + 3x, x)$.

$$\begin{aligned} x &= a(2+3x)+bx \Rightarrow \\ x &= 2a+(3a+b)x \Rightarrow \\ a &= 0 \text{ y } b = 1; \end{aligned}$$

por tanto, $x \in \text{gn}(2+3x, x)$. Así que $\text{gn}(1,x) \subset \text{gn}(2+3x,x)$. Por otra parte,

$$\begin{aligned} 2+3x &= 2\cdot 1+3\cdot x, \\ x &= 0\cdot 1+1\cdot x, \end{aligned}$$

por lo que $2+3x, x \in \text{gn}(1,x)$; luego $\text{gn}(2+3x,x) \subset \text{gn}(1,x)$. Por tanto,

$$\text{gn}(2+3x,x) = \text{gn}(1,x). \quad \blacksquare$$

Dependencia e independencia lineal

35 Sean $f_1(x) = 1$, $f_2(x) = e^{ax}$ y $f_3(x) = e^{bx}$, con $a \neq 0$, $b \neq 0$ y $a \neq b$; y $J = \mathbb{R}$.

(a) Probar que las funciones f_1, f_2 y f_3 son L.I. en J; es decir, en $\mathscr{F}(J)$, aplicando directamente la definición 3.15 (pág. 151).

(b) Demostrar que las funciones f_1, f_2 y f_3 son L.I. en J utilizando el wronskiano de estas funciones (cfr. ejemplo 3.44, pág. 155).

Solución *(a)* Sean $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ tales que

$$\alpha_1 \cdot 1 + \alpha_2 \cdot e^{ax} + \alpha_3 \cdot e^{bx} = 0 \quad \forall x \in \mathbb{R};$$

esto es,

$$\alpha_1 + \alpha_2 e^{ax} + \alpha_3 e^{bx} = 0 \quad \forall x \in \mathbb{R}.$$

En particular, si $x = 0$, $x = 1$ y $x = 2$ se tiene, respectivamente,

$$\begin{aligned} \alpha_1+\alpha_2+\alpha_3 &= 0, \\ \alpha_1+\alpha_2 e^{a}+\alpha_3 e^{b} &= 0 \quad \text{y} \\ \alpha_1+\alpha_2 e^{2a}+\alpha_3 e^{2b} &= 0. \end{aligned}$$

Y puesto que

$$\begin{aligned} \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^a & e^b \\ 1 & e^{2a} & e^{2b} \end{bmatrix} &\sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & e^a-1 & e^b-1 \\ 0 & e^{2a}-1 & e^{2b}-1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & \frac{e^b-1}{e^a-1} \\ 0 & e^{2a}-1 & e^{2b}-1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & \frac{e^b-1}{e^a-1} \\ 0 & 0 & (1-e^b)(e^a-e^b) \end{bmatrix}, \end{aligned}$$

y $a \neq 0$, $b \neq 0$, $a \neq b$, se debe tener $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Por lo que f_1, f_2 y f_3 son L.I. en J.

(b) Para cada $x \in J$,

$$W(f_1, f_2, f_3)(x) = \begin{vmatrix} 1 & e^{ax} & e^{bx} \\ 0 & ae^{ax} & be^{bx} \\ 0 & a^2e^{ax} & b^2e^{bx} \end{vmatrix}.$$

En particular,

$$\begin{aligned} W(f_1, f_2, f_3)(0) &= \begin{vmatrix} 1 & 1 & 1 \\ 0 & a & b \\ 0 & a^2 & b^2 \end{vmatrix} \\ &= ab^2 - ba^2 \\ &= ab(b-a) \\ &\neq 0 \end{aligned}$$

pues $a \neq b$ y $a \neq 0, b \neq 0$. Luego f_1, f_2 y f_3 son L.I. en J. ✔

36 Sean

$$A = \begin{bmatrix} -1 & 2 & 1 \\ 2 & -3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 3 & 1 \\ -1 & 1 & 1 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} -4 & 1 & 1 \\ 5 & -7 & 1 \end{bmatrix}.$$

Determinar si A, B y C son L.I. o L.D. en $\mathfrak{M}_{2\times 3}$.

Solución Las matrices son L.D. si el sistema

$$\alpha_1 A + \alpha_2 B + \alpha_3 C = \mathscr{O}$$

tiene solución no trivial; el cual equivale a

$$\begin{bmatrix} -\alpha_1 + 2\alpha_2 - 4\alpha_3 & 2\alpha_1 + 3\alpha_2 + \alpha_3 & \alpha_1 + \alpha_2 + \alpha_3 \\ 2\alpha_1 - \alpha_2 + 5\alpha_3 & -3\alpha_1 + \alpha_2 - 7\alpha_3 & \alpha_1 + \alpha_2 + \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

es decir,

$$\begin{aligned} -\alpha_1 + 2\alpha_2 - 4\alpha_3 &= 0 \\ 2\alpha_1 + 3\alpha_2 + \alpha_3 &= 0 \\ \alpha_1 + \alpha_2 + \alpha_3 &= 0 \\ 2\alpha_1 - \alpha_2 + 5\alpha_3 &= 0 \\ -3\alpha_1 + \alpha_2 - 7\alpha_3 &= 0 \\ \alpha_1 + \alpha_2 + \alpha_3 &= 0. \end{aligned}$$

Y como

$$\begin{bmatrix} -1 & 2 & -4 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \\ 2 & -1 & 5 \\ -3 & 1 & -7 \\ 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} -1 & 2 & -4 \\ 0 & 7 & -7 \\ 0 & 3 & -3 \\ 0 & 3 & -3 \\ 0 & -5 & 5 \\ 0 & 3 & -3 \end{bmatrix} \sim \begin{bmatrix} -1 & 2 & -4 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

se tiene

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} -2r \\ r \\ r \end{bmatrix} = r \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}.$$

Así, el sistema $\alpha_1 A + \alpha_2 B + \alpha_3 C = \mathscr{O}$ tiene solución no trivial (por ejemplo $\alpha_1 = -2, \alpha_2 = 1 = \alpha_3$); por tanto, las matrices A, B y C son L.D. en $\mathfrak{M}_{2\times 3}$. ✔

37 Determinar si los polinomios $p = x^3 - 4x^2 + 2x + 3$, $q = x^3 + 2x^2 + 4x - 1$ y $r = 2x^3 - x^2 + 3x + 5$ son L.I. o L.D. en $\mathbf{P}_3$.

Solución Se requiere encontrar escalares α_1, α_2 y α_3 tales que

$$\alpha_1 p + \alpha_2 q + \alpha_3 r = 0,$$

donde 0 es el polinomio constante cero; esto es,

$$(\alpha_1 + \alpha_2 + 2\alpha_3)x^3 + (-4\alpha_1 + 2\alpha_2 - \alpha_3)x^2 + (2\alpha_1 + 4\alpha_2 + 3\alpha_3)x + 3\alpha_1 - \alpha_2 + 5\alpha_3 = 0$$

que equivale, por igualdad de polinomios, a

$$\begin{aligned} \alpha_1 + \alpha_2 + 2\alpha_3 &= 0 \\ -4\alpha_1 + 2\alpha_2 - \alpha_3 &= 0 \\ 2\alpha_1 + 4\alpha_2 + 3\alpha_3 &= 0 \\ 3\alpha_1 - \alpha_2 + 5\alpha_3 &= 0. \end{aligned}$$

Puesto que

$$\begin{aligned} \begin{bmatrix} 1 & 1 & 2 \\ -4 & 2 & -1 \\ 2 & 4 & 3 \\ 3 & -1 & 5 \end{bmatrix} &\sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 6 & 7 \\ 0 & 2 & -1 \\ 0 & -4 & -1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & -1 \\ 0 & 6 & 7 \\ 0 & -4 & -1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 10 \\ 0 & 0 & -3 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 10 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

se tiene que $\alpha_1 p + \alpha_2 q + \alpha_3 r = 0$ tiene por solución $\alpha_1 = \alpha_2 = \alpha_3 = 0$, por lo que p, q y r son L.I. en $\mathbf{P}_3$. ✔

38 Sean $\vec{u}_1$, $\vec{u}_2$ y $\vec{u}_3$ vectores L.I. en un espacio vectorial $\mathbf{E}$. Probar que $\vec{v}_1 = 5\vec{u}_1$, $\vec{v}_2 = 3\vec{u}_1 - \vec{u}_2$ y $\vec{v}_3 = \vec{u}_1 + \vec{u}_3$ son también linealmente independientes.

DEMOSTRACIÓN ■ Sean $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ tales que

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \alpha_3\vec{v}_3 = \vec{0}_{\mathbf{E}},$$

entonces

$$(5\alpha_1 + 3\alpha_2 + \alpha_3)\vec{u}_1 + (-\alpha_2)\vec{u}_2 + \alpha_3\vec{u}_3 = \vec{0}_{\mathbf{E}},$$

y puesto que $\vec{u}_1$, $\vec{u}_2$ y $\vec{u}_3$ son L.I.,

$$\begin{aligned} 5\alpha_1 + 3\alpha_2 + \alpha_3 &= 0 \\ -\alpha_2 &= 0 \\ \alpha_3 &= 0, \end{aligned}$$

de donde $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Luego $\vec{u}_1$, $\vec{u}_2$ y $\vec{u}_3$ son L.I. ■

39 Sean $\vec{u}_1, \vec{u}_2 \in \mathbb{R}^n$ vectores L.I. y $A \in \mathfrak{M}_n$ una matriz no singular (invertible). Demostrar que $\vec{v}_1 = A\vec{u}_1$ y $\vec{v}_2 = A\vec{u}_2$ son también vectores L.I.

DEMOSTRACIÓN ■ Sean α_1 y α_2 escalares tales que

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 = \vec{0}_{\mathbb{R}^n},$$

esto es,

$$\begin{aligned} \vec{0}_{\mathbb{R}^n} &= \alpha_1(A\vec{u}_1) + \alpha_2(A\vec{u}_2) \\ &= A(\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2). \end{aligned}$$

Como A es no singular, la anterior igualdad implica

$$\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 = A^{-1}\vec{0}_{\mathbb{R}^n} = \vec{0}_{\mathbb{R}^n},$$

y dado que $\vec{u}_1$ y $\vec{u}_2$ son vectores L.I., se debe tener $\alpha_1 = \alpha_2 = 0$; por tanto, $\vec{v}_1$ y $\vec{v}_2$ son L.I. ■

40 Sean $\mathbf{E}$ un espacio vectorial, $\vec{u}_1, \ldots, \vec{u}_k$ vectores en este espacio, y $S = \text{gn}(\vec{u}_1, \ldots, \vec{u}_k)$. Demostrar que los vectores $\vec{u}_i$ son L.I. si y sólo si todo vector $\vec{u} \in S$ se puede escribir como combinación lineal de los vectores $\vec{u}_i$ de manera unívoca (esto es, si $\vec{u} = \sum_{i=1}^k a_i u_i$ y $\vec{u} = \sum_{i=1}^k b_i\vec{u}_i$, entonces $a_i = b_i \forall i$).

DEMOSTRACIÓN ■ ($\Rightarrow$) Si los vectores $\vec{u}_i$ son L.I. y $\vec{u} = \sum_{i=1}^k a_i\vec{u}_i = \sum_{i=1}^k b_i\vec{u}_i$, entonces

$$\begin{aligned} \vec{0}_{\mathbf{E}} &= \textstyle\sum_{i=1}^k a_i\vec{u}_i - \sum_{i=1}^k a_i\vec{u}_i \\ &= \textstyle\sum_{i=1}^k (a_i - b_i)\vec{u}_i \end{aligned}$$

implica $a_i = b_i \; \forall i = 1, \ldots, k$. ($\Leftarrow$) Sean a_i escalares tales que $\sum_{i=1}^k a_i\vec{u}_i = \vec{0}_{\mathbf{E}}$. Si todo vector de S se puede escribir de manera unívoca como combinación lineal de los vectores $\vec{u}_i$, en particular $\vec{0}_{\mathbf{E}} = \sum_{i=1}^k 0\vec{u}_i$; por tanto, $a_i = 0$ para todo $i = 1, \ldots, k$. Lo cual prueba que los vectores $\vec{u}_i$ son L.I. ■

41 Probar que las funciones $f_1(x) = \cos x$, $f_2(x) = \cos 2x$, $f_3(x) = \cos 3x$ y $f_4(x) = \cos 4x$ son L.I. en $J = (-\infty, \infty)$; es decir, en $\mathscr{F}(J)$.

DEMOSTRACIÓN ■

$$W(f_1, f_2 f_3, f_4)(x) = \begin{vmatrix} \cos x & \cos 2x & \cos 3x & \cos 4x \\ -\operatorname{sen} x & -2\operatorname{sen} 2x & -3\operatorname{sen} 3x & -4\operatorname{sen} 4x \\ -\cos x & -2^2\cos 2x & -3^2\cos 3x & -4^2\cos 4x \\ \operatorname{sen} x & 2^3\operatorname{sen} 2x & 3^3\operatorname{sen} 3x & 4^3\operatorname{sen} 4x \end{vmatrix}.$$

$$\begin{aligned} W(f_1, f_2 f_3, f_4)(\pi/4) &= \begin{vmatrix} \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} & -1 \\ -\frac{\sqrt{2}}{2} & -2 & -3\frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & 0 & 3^2\frac{\sqrt{2}}{2} & 4^2 \\ \frac{\sqrt{2}}{2} & 2^3 & 3^3\frac{\sqrt{2}}{2} & 0 \end{vmatrix} \\ &= \left(\frac{\sqrt{2}}{2}\right)^2 \begin{vmatrix} 1 & 0 & -1 & -1 \\ -1 & -2 & -3 & 0 \\ -1 & 0 & 3^2 & 4^2 \\ 1 & 2^3 & 3^3 & 0 \end{vmatrix} \\ &= \left(\frac{\sqrt{2}}{2}\right)^2 \begin{vmatrix} 1 & 0 & -1 & -1 \\ 0 & -2 & -4 & -1 \\ 0 & 0 & 8 & 15 \\ 0 & 8 & 28 & 1 \end{vmatrix} \\ &= \left(\frac{\sqrt{2}}{2}\right)^2 \begin{vmatrix} 1 & 0 & -1 & -1 \\ 0 & -2 & 2 & -1 \\ 0 & 0 & 8 & 15 \\ 0 & 0 & 12 & -3 \end{vmatrix} \\ &= \left(\frac{\sqrt{2}}{2}\right)^2 (-2)(-24-180) \\ &= 204 \neq 0. \end{aligned}$$

Ya que el wronskiano de estas funciones es distinto de cero en un punto de J (cfr. ejemplo 3.44, pág. 155), f_1, f_2, f_3 y f_4 son L.I. en $J = (-\infty, \infty)$. ■

Bases y dimensión

42 Demostrar que el conjunto infinito $U = \{e^x, e^{2x}, \ldots, e^{nx}, \ldots\}$ es L.I. en $\mathscr{F}(\mathbb{R})$, el espacio de funciones $f : \mathbb{R} \to \mathbb{R}$.

DEMOSTRACIÓN ■ Sea F un subconjunto finito cualquiera de U, entonces F está contenido en un conjunto de la forma $F_n = \{e^x, e^{2x}, \ldots, e^{nx}\}$ para algún entero positivo n; luego es suficiente probar que este último conjunto es L.I. para que F sea L.I.; por tanto, U también. Se procede por inducción sobre n: si $n = 1$, entonces $F_n = \{e^x\}$ que es L.I. porque $f(x) = e^x$ no es la función constante cero. Sea n un entero mayor a 1 y suponga que el conjunto $F_{n-1} = \{e^x, e^{2x}, \ldots, e^{(n-1)x}\}$ es L.I. Sean α_i, $i = 1, 2, \ldots, n$, escalares tales que

$$\alpha_1 e^x + \alpha_2 e^{2x} + \cdots + \alpha_{n-1} e^{(n-1)x} + \alpha_n e^{nx} = 0 \qquad \text{(L1)}$$

para todo $x \in \mathbb{R}$. Al derivar la precedente igualdad se obtiene

$$\alpha_1 e^x + 2\alpha_2 e^{2x} + \cdots + (n-1)\alpha_{n-1} e^{(n-1)x} + n\alpha_n e^{nx} = 0 \qquad \text{(L2)}$$

para todo $x \in \mathbb{R}$. Al multiplicar la igualdad (L1) por n y restar con la igualdad (L2) se produce

$$\alpha_1(n-1)e^x + \alpha_2(n-2)e^{2x} + \cdots + \alpha_{n-1}(n-(n-1))e^{(n-1)x} = 0$$

para todo $x \in \mathbb{R}$. Por hipótesis de inducción las funciones e^x, $e^{2x}, \ldots$, $e^{(n-1)x}$ son L.I., y puesto que $n \neq 1, \ldots, n-1$, la igualdad anterior implica

$$\alpha_1 = \alpha_2 = \cdots = \alpha_{n-1} = 0.$$

Luego, al sustituir estos valores en (L1) se desprende que

$$\alpha_n e^{nx} = 0 \;\; \forall x \in \mathbb{R};$$

y por tanto (ya que $f(x) = e^{nx}$ no es la función constante cero) $\alpha_n = 0$. Lo cual prueba que las funciones e^{kx}, $k = 1, \ldots, n$, son L.I. en $\mathscr{F}(\mathbb{R})$ para todo n. ■

43 Sea $\mathbf{P}_n$ el espacio de polinomios de grado a lo más n y

$$S = \{p \in \mathbf{P}_n \,|\, p(0) = 0\}.$$

Mostrar que $S < \mathbf{P}_n$ y determinar $\dim(S)$.

Solución (1) Claramente el polinomio constante cero, θ, satisface $\theta(0) = 0$; por lo que $\theta \in S$. (2) Si $p_1, p_2 \in S$, entonces $p_1(0) = p_2(0) = 0$; por lo que $p_1(0) + p_2(0) = 0$; por tanto, $p_1 + p_2 \in S$. (3) Si $p \in S$ y $\alpha \in \mathbb{R}$, entonces $\alpha p(0) = \alpha 0 = 0$; por tanto, $\alpha p \in S$. De (1), (2) y (3), $S < P_n$. Los polinomios x, $x^2, \ldots, x^n$ pertenecen a S y son L.I. (pues forman parte del conjunto L.I. $1, x, \ldots, x^n$). Si $p \in S$, $x = 0$ es raíz de p; por tanto, existe un polinomio q de grado a lo más $n-1$, $q = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$, tal que

$$\begin{aligned} p &= xq \\ &= a_0 x + a_1 x^2 + \cdots + a_{n-1} x^n. \end{aligned}$$

Luego los polinomios L.I. x, $x^2, \ldots, x^n$ generan a S, por tanto forman una base para este subespacio; así que

$$\dim(S) = n. \quad ✔$$

44 Sean $\alpha, \beta \in \mathbb{R}$ un par de números reales con $\beta \neq 0$. Sea S el conjunto de sucesiones de números reales que satisfacen la relación

$$x_{n+2} + \alpha x_{n+1} + \beta x_n = 0 \; \forall n \qquad (*)$$

(a) Mostrar que $S < \mathbb{R}^\infty$.

(b) Probar que si $\lambda, \mu \in \mathbb{R}$ son raíces distintas del polinomio $X^2 + \alpha X + \beta = 0$, entonces las sucesiones $u = (u_n)$ y $v = (v_n)$, con $u_n = \lambda^n$ y $v_n = \mu^n$, pertenecen a S.

(c) Demostrar que $\{u, v\}$ es una base de S.

DEMOSTRACIÓN ■ *(a)* Claramente la sucesión constante cero satisface la relación (*). Sean $x = (x_n)$ y $y = (y_n)$ un par de sucesiones en S y $a, b \in \mathbb{R}$, entonces,

$$\begin{aligned} ax_{n+2} + by_{n+2} + \alpha(ax_{n+1} + by_{n+1}) + \beta(ax_n + by_n) &= a(x_{n+2} + \alpha x_{n+1} + \beta x_n) \\ &\quad + b(y_{n+2} + \alpha y_{n+1} + \beta y_n) \\ &= a0 + b0 \\ &= 0; \end{aligned}$$

por tanto, $ax + by \in S$. Lo cual prueba $S < \mathbb{R}^\infty$.

(b) Puesto que λ es raíz de $X^2 + \alpha X + \beta = 0$, se tiene

$$\begin{aligned} u_{n+2} + \alpha u_{n+1} + \beta u_n &= \lambda^{n+2} + \alpha \lambda^{n+1} + \beta \lambda^n \\ &= \lambda^n(\lambda^2 + \alpha\lambda + \beta) \\ &= \lambda^n \cdot 0 \\ &= 0 \; \forall n. \end{aligned}$$

Análogamente se prueba que $v_{n+2} + \alpha v_{n+1} + \beta v_n = 0$. Así $u, v \in S$.

(c) *(i)* Dado que $\beta \neq 0$, se concluye $\lambda \neq 0$, $\mu \neq 0$. Sean $a, b \in \mathbb{R}$ tales que $au + bv = 0$. Entonces

$$\begin{aligned} a\lambda + b\mu &= 0 \quad \text{y} \\ a\lambda^2 + b\mu^2 &= 0. \end{aligned}$$

Al multiplicar la primer igualdad por $-\mu$ y sumarla con la segunda se obtiene

$$a\lambda(\lambda - \mu) = 0,$$

y puesto que $\lambda \neq 0$, $\mu = 0$ y $\lambda \neq \mu$ se desprende que $a = 0$ y $b = 0$. Luego u y v son L.I.

(ii) Dado que λ y μ son raíces de la ecuación cuadrática $X^2 + \alpha X + \beta = 0$, satisfacen

$$\begin{aligned} -\alpha &= \lambda + \mu \text{ y} \\ \beta &= \lambda\mu. \end{aligned}$$

Sea $x = (x_n) \in S$. Sean $C_1, C_2 \in \mathbb{R}$ las únicas soluciones del sistema

$$\begin{aligned} C_1\lambda + C_2\mu &= x_1 \\ C_1\lambda^2 + C_2\mu^2 &= x_2; \end{aligned}$$

esto es,

$$\begin{aligned} C_1 &= \frac{\mu x_1 - x_2}{\lambda(\mu - \lambda)} \\ C_2 &= \frac{x_2 - \lambda x_1}{\mu(\mu - \lambda)} \end{aligned}$$

Por construcción de C_1 y C_2,

$$x_n = C_1\lambda^n + C_2\mu^n \qquad (**)$$

se cumple para $n = 1$ y $n = 2$. Sea k un entero mayor a 2 y suponga que la precedente igualdad es válida para todo entero m con $2 < m \leq k-1$. Entonces, por (*),

$$\begin{aligned} x_k &= -\alpha x_{k-1} - \beta x_{k-2} \\ &= (\lambda+\mu)x_{k-1} - \lambda\mu x_{k-2} \\ &= (\lambda+\mu)(C_1\lambda^{k-1} + C_2\mu^{k-1}) - \lambda\mu(C_1\lambda^{k-2} + C_2\mu^{k-2}) \\ &= C_1\lambda^k + C_2\mu^k + C_1\lambda^{k-1}\mu + C_2\mu^{k-1}\lambda - C_1\lambda^{k-1}\mu - C_2\mu^{k-1}\lambda \\ &= C_1\lambda^k + C_2\mu^k. \end{aligned}$$

Se sigue por inducción que (**) se cumple para todo n; por tanto, $x = C_1u + C_2v$ para todo $x \in S$; es decir, u, v generan a S. ■

45 Sean R, S un par de subespacios de un espacio vectorial $\mathbf{E}$ de dimensiones finitas r y s, respectivamente. Mostrar que

$$\dim(R+S) = \dim(R) + \dim(S) - \dim(R \cap S),$$

donde $R+S$ es la suma de estos subespacios definida en el ejercicio resuelto 26 de esta sección y $R \cap S$ es el subespacio formado por la intersección de los subespacios R y S (cfr. ejercicio resuelto 30).

DEMOSTRACIÓN ■ Sean $\{\vec{u}_1, \ldots, \vec{u}_k\}$ una base de $R \cap S$. Puesto que $R \cap S < R$ y $R \cap S < S$, se puede completar, mediante el proceso dado en la página 167, ésta a una base $\{\vec{u}_1, \ldots, \vec{u}_k, \vec{e}_1, \ldots, \vec{e}_{r-k}\}$ de R y a una base $\{\vec{u}_1, \ldots, \vec{u}_k, \vec{f}_1, \ldots, \vec{f}_{s-k}\}$ de S. Sea $\vec{u} + \vec{v} \in R+S$, entonces

$$\vec{u} = \sum_{i=1}^{k} \alpha_i\vec{u}_i + \sum_{i=1}^{r-k} \beta_i\vec{e}_i$$

para algunos escalares α_i, β_i, y

$$\vec{v} = \sum_{i=1}^{k} \alpha_i'\vec{u}_i + \sum_{i=1}^{s-k} \gamma_i\vec{f}_i$$

para algunos escalares α_i', γ_i; luego

$$\vec{u} + \vec{v} = \sum_{i=1}^{k} (\alpha_i + \alpha_i')\vec{u}_i + \sum_{i=1}^{r-k} \beta_i\vec{e}_i + \sum_{i=1}^{s-k} \gamma_i\vec{f}_i$$

lo cual prueba que $R+S$ está generado por los $r+s-k$ vectores $\vec{u}_1, \ldots, \vec{u}_k$, $\vec{e}_1, \ldots, \vec{e}_{r-k}$, $\vec{f}_1, \ldots, \vec{f}_{s-k}$. Sean α_i, β_i, γ_i escalares tales que

$$\sum_{i=1}^{k}\alpha_i\vec{u}_i + \sum_{i=1}^{r-k}\beta_i\vec{e}_i + \sum_{i=1}^{s-k}\gamma_i\vec{f}_i = \vec{0}_{\mathbf{E}}.$$

De donde

$$\sum_{i=1}^{s-k}\gamma_i\vec{f}_i = -\sum_{i=1}^{k}\alpha_i\vec{u}_i - \sum_{i=1}^{r-k}\beta_i\vec{e}_i \in R;$$

por tanto,

$$\sum_{i=1}^{s-k} \gamma_i \vec{f}_i \in R \cap S.$$

Así que existen escalares $-\delta_i$ tales que

$$\sum_{i=1}^{s-k} \gamma_i \vec{f}_i = \sum_{i=1}^{k} (-\delta_i) \vec{u}_i;$$

por ende,

$$\sum_{i=1}^{k} \delta_i \vec{u}_i + \sum_{i=1}^{s-k} \gamma_i \vec{f}_i = \vec{0}_{\mathbf{E}}.$$

Como los vectores $u_1, \ldots, \vec{u}_k, \vec{f}_1, \ldots, \vec{f}_{s-k}$ son L.I., se debe tener entonces que los escalares γ_i son nulos; por tanto

$$\sum_{i=1}^{k} \alpha_i \vec{u}_i + \sum_{i=1}^{r-k} \beta_i \vec{e}_i = \vec{0}_{\mathbf{E}}$$

y ya que los vectores $\vec{u}_1, \ldots, \vec{u}_k, \vec{e}_1, \ldots, \vec{e}_{r-k}$ son L.I., se tiene que los escalares α_i, β_i son también nulos. Se ha probado así que los vectores $\vec{u}_1, \ldots, \vec{u}_k$, $\vec{e}_1, \ldots, \vec{e}_{r-k}$, $\vec{f}_1, \ldots, \vec{f}_{s-k}$ forman una base de $R + S$, por lo que

$$\begin{aligned} \dim(R+S) &= r+s-k \\ &= \dim(R) + \dim(S) - \dim(R \cap S). \quad \blacksquare \end{aligned}$$

46 (**Ecuación de un subespacio de** $\mathbb{R}^n$). Sea S el subespacio generado por $(1,2,1)$ y $(-1,2,2)$. Encontrar un sistema homogéneo $A\vec{x} = \vec{0}_{\mathbb{R}^3}$ tal que su espacio solución (espacio nulo) sea el espacio S. A dicho sistema, $A\vec{x} = \vec{0}_{\mathbb{R}^3}$, se le llama ecuación del subespacio S.

Solución **Un primer método:** Si $(x,y,z) \in S$, entonces

$$\begin{aligned} (x,y,z) &= a(1,2,1) + b(-1,2,2) \\ &= (a-b, 2a+2b, a+2b); \end{aligned}$$

por tanto, se debe tener

$$\begin{aligned} a - b &= x \\ 2a + 2b &= y \\ a + 2b &= z; \end{aligned}$$

que implican

$$\begin{aligned} y - 2x &= 4b \\ z - x &= 3b; \end{aligned}$$

por ende,

$$2x - 3y + 4z = 0$$

es la ecuación del subespacio S.

Un segundo método: Si $(x,y,z) \in S$, entonces (x,y,z) es combinación lineal del par de vectores L.I. $(1,2,1)$ y $(-1,2,2)$. Así que si A es la matriz cuyas filas son estos vectores, al llevar a forma escalonada la última fila debe ser nula:

$$\begin{bmatrix} 1 & 2 & 1 \\ -1 & 2 & 2 \\ x & y & z \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 4 & 3 \\ 0 & y-2x & z-x \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 4 & 3 \\ 0 & 0 & 2x-3y+4z \end{bmatrix};$$

por tanto,

$$2x-3y+4z=0$$

es la ecuación del subespacio S. ✔

47 Sean $S_1 = \text{gn}((1,0,-1),(1,-1,0))$ y $S_2 = \text{gn}((2,1,0),(-2,0,1))$. Hallar:

(a) Una base y la dimensión de $S_1 + S_2$.

(b) Una base y la dimensión de $S_1 \cap S_2$.

Solución *(a)* $S_1 + S_2$ está generado por $(1,0,-1),(1,-1,0),(2,1,0),(-2,0,1)$ y puesto que

$$\begin{bmatrix} 1 & 1 & 2 & -2 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & -2 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 3 & -1 \end{bmatrix}$$

una base para este espacio es $\{(1,0,-1),(1,-1,0),(2,1,0)\}$. Por tanto, $\dim(S_1+S_2)=3$; es decir, $S_1+S_2=\mathbb{R}^3$.

(b) Dado que

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \\ x & y & z \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & y & x+z \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & x+y+z \end{bmatrix};$$

una ecuación para el espacio S_1 es

$$x+y+z=0.$$

Puesto que

$$\begin{bmatrix} 2 & 1 & 0 \\ -2 & 0 & 1 \\ x & y & z \end{bmatrix} \sim \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 2y-x & 2z \end{bmatrix}$$
$$\sim \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & x-2y+2z \end{bmatrix};$$

una ecuación para el subespacio S_2 es

$$x-2y+2z=0.$$

Entonces los vectores (x,y,z) que pertenecen a $S_1\cap S_2$ son las soluciones del sistema homogéneo

$$\begin{aligned} x+y+z &= 0\\ x-2y+2z &= 0. \end{aligned}$$

Y ya que

$$\begin{bmatrix} 1 & 1 & 1\\ 1 & -2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1\\ 0 & -3 & 1 \end{bmatrix},$$

se tiene

$$\begin{bmatrix} x\\ y\\ z \end{bmatrix} = \begin{bmatrix} -\frac{4}{3}r\\ \frac{1}{3}r\\ r \end{bmatrix} = r\begin{bmatrix} -\frac{4}{3}\\ \frac{1}{3}\\ 1 \end{bmatrix}.$$

Por lo que una base de $S_1\cap S_2$ es $\{(-4,1,3)\}$ y $\dim(S_1\cap S_2)=1$. ✔

En los ejercicios 48 a 51, X es un conjunto no vacío, y $\mathscr{F}(X)$ es el espacio de funciones con dominio en X y valores en $\mathbb{R}$; si $A\subset X$ es un subconjunto de X, se definen

$$\chi_A(x)=\begin{cases} 1, \text{ si } x\in A;\\ 0, \text{ si } x\notin A. \end{cases}$$

(La función χ_A se llama función característica o indicadora de A) y

$$F_A=\{f\in\mathscr{F}(X)\,|\,f(x)=0\ \forall x\in X-A\}.$$

48 Probar que F_A es un subespacio de $\mathscr{F}(X)$.

DEMOSTRACIÓN ■ (1) La función constante cero, θ, satisface $\theta(x)=0$ para todo $x\in X$, en particular es nula en todo $x\in X-A$; luego $\theta\in F_A$. (2) Si α,β son un par de escalares y $f,g\in F_A$, entonces $f(x)=g(x)=0$ para todo $x\in X-A$. Por tanto,

$$(\alpha f+\beta g)(x)=\alpha f(x)+\beta g(x)=0\quad \forall x\in X-A;$$

lo cual prueba que $\alpha f+\beta g\in F_A$. De (1) y (2) se concluye que $F_A<\mathscr{F}(X)$. ■

49 Si $A_1,\ldots,A_k$ forman una partición de X; es decir, $X=A_1\cup\cdots\cup A_k$ y $A_i\cap A_j=\emptyset$ si $i\neq j$, demostrar que

$$\mathscr{F}(X)=F_{A_1}\oplus\cdots\oplus F_{A_k};$$

esto es, $\mathscr{F}(X)$ es la suma directa (cfr. el ejercicio resuelto 27 de esta sección) de los subespacios F_{A_i}.

DEMOSTRACIÓN ■ Sea $f\in\mathscr{F}(X)$, entonces

$$f=\chi_{A_1}f+\cdots+\chi_{A_k}f.$$

En efecto, si $x \in X$, existe un único r tal que $x \in A_r$, luego

$$(\chi_{A_r} f)(x) = \chi_{A_r}(x) f(x) = 1 \cdot f(x) = f(x)$$

y

$$(\chi_{A_i} f)(x) = \chi_{A_i}(x) f(x) = 0 \cdot f(x) = 0 \quad \text{si } i \neq r;$$

lo cual prueba que $f = \sum_{i=1}^{k} \chi_{A_i} f$. Por otra parte, si $x \notin A_i$, entonces

$$(\chi_{A_i} f)(x) = \chi_{A_i}(x) f(x) = 0 \cdot f(x) = 0,$$

lo que demuestra que $\chi_{A_i} f \in F_{A_i}$. Supongamos que $f_1 + \cdots + f_k = \theta$, la función constante cero, con $f_i \in F_{A_i}$. Sea $x \in X$ y r el único subíndice tal que $x \in A_r$; entonces $f_j(x) = 0\ \ \forall\, j = r$. Luego $f_r(x) = f_1(x) + \cdots + f_r(x) + \cdots + f_k(x) = 0$ y, por tanto, $f_r(x) = 0$. Lo cual prueba $f_r = \theta\ \forall\, r = 1, \ldots, k$. Así, toda función $f \in \mathscr{F}(X)$ se puede escribir en la forma $f = f_1 + \cdots + f_k$, con $f_i \in F_{A_i}$, de manera única. ■

50 Si los conjuntos A_i, $i = 1, \ldots, k$, forman una partición de X, probar que las funciones χ_{A_i} son L.I.

DEMOSTRACIÓN ■ Sean los escalares α_i tales que

$$\alpha_1 \chi_{A_1}(x) + \cdots + \alpha_k \chi_{A_k}(x) = 0 \quad \forall x \in X.$$

Sea $x_i \in A_i$, entonces

$$\begin{aligned} 0 &= \alpha_1 \chi_{A_1}(x) + \cdots + \alpha_k \chi_{A_k}(x) \\ &= \alpha_i. \end{aligned}$$

Luego las funciones χ_{A_i} son L.I. ■

51 Si X es un conjunto finito, $X = \{x_1, \ldots, x_k\}$, hallar una base y la dimensión de $\mathscr{F}(X)$.

Solución Sea $f \in \mathscr{F}(X)$ y $c_i = f(x_i)$ para cada i. Por los ejercicios precedentes, ya que los conjuntos $\{x_i\}$ forman una partición de X, las funciones $\chi_{\{x_i\}}$ son L.I., y

$$f = \chi_{\{x_1\}} f + \cdots + \chi_{\{x_k\}} f;$$

es decir,

$$f = c_1 \chi_{\{x_1\}} + \cdots + c_k \chi_{\{x_k\}}.$$

Por lo que $\{\chi_{\{x_1\}}, \ldots, \chi_{\{x_k\}}\}$ es una base de $\mathscr{F}(X)$; por tanto,

$$\dim(\mathscr{F}(X)) = k. \quad ✔$$

Espacios vectoriales complejos

52 Sea $S = \{\vec{z} = (z_1, z_2, z_3, z_4) \in \mathbb{C}^4 \mid z_1 = (1+i)z_2 - iz_3 = z_2 - z_4 = 0\}$. Mostrar que $S < \mathbb{C}^4$ y encontrar una base y la dimensión de S.

Solución Los vectores $\vec{z}=(z_1,z_2,z_3,z_4)\in S$ satisfacen las relaciones $z_1=0$, $z_2=\frac{i}{1+i}z_3=(\frac{1}{2}+\frac{1}{2}i)z_3$ y $z_2=z_4$; es decir,

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} 0 \\ (\frac{1}{2}+\frac{1}{2}i)r \\ r \\ (\frac{1}{2}+\frac{1}{2}i)r \end{bmatrix} = r\begin{bmatrix} 0 \\ (\frac{1}{2}+\frac{1}{2}i) \\ 1 \\ (\frac{1}{2}+\frac{1}{2}i) \end{bmatrix}; \quad r\in\mathbb{C}.$$

Con lo que $S=\operatorname{gn}((0,\frac{1}{2}+\frac{1}{2}i,1,\frac{1}{2}+\frac{1}{2}i))$; luego S es un subespacio de $\mathbb{C}^4$, una base de S es

$$\{(0,\tfrac{1}{2}+\tfrac{1}{2}i,1,\tfrac{1}{2}+\tfrac{1}{2}i)\} \text{ y } \dim(S)=1. \quad ✔$$

53 Determinar si los vectores $\vec{z}_1=(1+i,2-i,1)$, $\vec{z}_2=(3+2i,2-2i,1)$ y $\vec{z}_3=(1,2,-i)$ son L.I. o L.D. en $\mathbb{C}^3$.

Solución Dado que

$$\begin{aligned}\begin{bmatrix} 1+i & 3+2i & 1 \\ 2-i & 2-2i & 2 \\ 1 & 1 & -i \end{bmatrix} &\sim \begin{bmatrix} 1 & 1 & -i \\ 2-i & 2-2i & 2 \\ 1+i & 3+2i & 1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & -i \\ 0 & -i & 3+2i \\ 0 & 2+i & i \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & -i \\ 0 & 1 & -2+3i \\ 0 & 2+i & i \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & -i \\ 0 & 1 & -2+3i \\ 0 & 2+i & i \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & -i \\ 0 & 1 & -2+3i \\ 0 & 0 & 7-3i \end{bmatrix},\end{aligned}$$

el sistema $\alpha_1\vec{z}_1+\alpha_2\vec{z}_2+\alpha_3\vec{z}_3=(0,0,0)$, $\alpha_i\in\mathbb{C}$, tiene únicamente la solución trivial $\alpha_1=\alpha_2=\alpha_3=0$, los vectores $\vec{z}_1$, $\vec{z}_2$ y $\vec{z}_3$ son L.I. ✔

54 Sean $\vec{z}_1=(1-i,i,2i,-1-2i)$, $\vec{z}_2=(1,-1,2i,1+3i)$, $\vec{z}_3=(1+2i,-3-2i,2i,5+13i)$ y $\vec{z}_4=(-i,-i,2+3i,-1)$; encontrar una base del subespacio $S=\operatorname{gn}(\vec{z}_1,\vec{z}_2,\vec{z}_3,\vec{z}_4)$ en $\mathbb{C}^4$.

Solución Puesto que

$$\begin{aligned}\begin{bmatrix} 1-i & 1 & 1+2i & -i \\ i & -1 & -3-2i & -i \\ 2i & 2i & 2i & 2+3i \\ -1-2i & 1+3i & 5+13i & -1 \end{bmatrix} &\sim \begin{bmatrix} 1-i & 1 & 1+2i & -i \\ 1 & i & -2+3i & -1 \\ 2i & 2i & 2i & 2+3i \\ -1-2i & 1+3i & 5+13i & -1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & i & -2+3i & -1 \\ 1-i & 1 & 1+2i & -i \\ 2i & 2i & 2i & 2+3i \\ -1-2i & 1+3i & 5+13i & -1 \end{bmatrix}\end{aligned}$$

$$\sim \begin{bmatrix} 1 & i & -2+3i & -1 \\ 0 & -i & -3i & 1-2i \\ 0 & 2+2i & 6+6i & 2+5i \\ 0 & -1+4i & -3+12i & -2-2i \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & i & -2+3i & -1 \\ 0 & 1 & 3 & 2+i \\ 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 4-9i \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & i & -2+3i & -1 \\ 0 & 1 & 3 & 2+i \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 4-9i \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & i & -2+3i & -1 \\ 0 & 1 & 3 & 2+i \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

los vectores que corresponden a columna con pivote en la forma escalonada forman una base para S; es decir, $\{\vec{z}_1, \vec{z}_2, \vec{z}_4\}$. ✔

55 Sea $\mathfrak{M}_2(\mathbb{C})$ el espacio vectorial de matrices 2×2 con componentes complejas y

$$S = \{A = (a_{ij}) \,|\, a_{11} + a_{22} = 0\}.$$

Comprobar que S es un subespacio vectorial de $\mathfrak{M}_2(\mathbb{C})$, y encontrar una base y la dimensión de S.

Solución Dado que toda matriz $A \in S$ tiene la forma

$$\begin{aligned} A &= \begin{bmatrix} \alpha & \beta \\ \gamma & -\alpha \end{bmatrix} \\ &= \alpha \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \beta \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \end{aligned}$$

es evidente que

$$S = \text{gn}\left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right);$$

por lo que $S < \mathfrak{M}_2(\mathbb{C})$. Por otro lado, si $a, b, c \in \mathbb{C}$ son escalares tales que

$$a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

se tiene

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix};$$

de donde $a = b = c = 0$; por tanto, estas matrices son L.I. en $\mathfrak{M}_2(\mathbb{C})$. Luego estas matrices forman una base de S; por ende, $\dim(S) = 3$. ✔

56 Sea $\mathbb{C}^2$ el espacio vectorial de vectores cuyos componentes son números complejos.

(a) Probar que $\vec{z}_1 = (1-i, i)$ y $\vec{z}_2 = (2, -1+i)$ son L.D. en $\mathbb{C}^2$.

(b) Sea **E** el espacio de pares ordenados (w_1, w_2), con $w_1, w_2 \in \mathbb{C}$, con la suma usual de vectores en $\mathbb{C}^2$ pero con la multiplicación por escalares restringida a los números reales. Probar que $\vec{z}_1$ y $\vec{z}_2$ son L.I. en **E**.

DEMOSTRACIÓN *(a)* Sean $\alpha, \beta \in \mathbb{C}$ tales que $\alpha\vec{z}_1 + \beta\vec{z}_2 = 0$, entonces,

$$((1-i)\alpha + 2\beta, i\alpha + (-1+i)\beta) = (0, 0);$$

por tanto,

$$\begin{aligned} (1-i)\alpha + 2\beta &= 0, \\ i\alpha + (-1+i)\beta &= 0. \end{aligned}$$

Y puesto que

$$\begin{aligned} \begin{bmatrix} 1-i & 2 \\ i & -1+i \end{bmatrix} &\sim \begin{bmatrix} i & -1+i \\ 1-i & 2 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1+i \\ 1-i & 2 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1+i \\ 0 & 0 \end{bmatrix}, \end{aligned}$$

se tiene

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} -(1+i)r \\ r \end{bmatrix} = r\begin{bmatrix} -(1+i) \\ 1 \end{bmatrix}.$$

Por lo que $\vec{z}_1$ y $\vec{z}_2$ son L.D. en $\mathbb{C}^2$.

(b) Sean $\alpha, \beta \in \mathbb{R}$ tales que $\alpha\vec{z}_1 + \beta\vec{z}_2 = 0$, entonces

$$((1-i)\alpha + 2\beta, i\alpha + (-1+i)\beta) = (0, 0);$$

por tanto,

$$\begin{aligned} \alpha + 2\beta - i\alpha &= 0, \\ -\beta + (\alpha+\beta)i &= 0. \end{aligned}$$

De donde $\alpha = 0$ y $\beta = 0$; por lo que $\vec{z}_1$ y $\vec{z}_2$ son L.I. en **E**. ■

57 Dado que $\mathbb{C}$ es un espacio vectorial es fácil probar que $\mathscr{F}(X, \mathbb{C})$, el conjunto de funciones $f: X \to \mathbb{C}$, es un espacio vectorial complejo para todo conjunto no vacío X. Sean $X = \mathbb{R}$ y $f_1(x) = 2 + 2i\cos(x)$, $f_2(x) = \operatorname{sen}^2(x) + i\cos^2(x)$ y $f_3(x) = \cos^2(x) - i\operatorname{sen}^2(x)$. Probar que f_1, f_2 y f_3 son L.I. en $\mathscr{F}(X, \mathbb{C})$.

Solución Sean $\alpha, \beta, \gamma \in \mathbb{C}$ tales que $\alpha f_1(x) + \beta f_2(x) + \gamma f_3(x) = 0$ para todo $x \in \mathbb{R}$, entonces

$$\begin{aligned} 2\alpha + \beta\operatorname{sen}^2(x) + \gamma\cos^2(x) &= 0, \\ 2\alpha\cos(x) + \beta\cos^2(x) - \gamma\operatorname{sen}^2(x) &= 0 \end{aligned}$$

para todo $x \in \mathbb{R}$. En particular, si $x = 0$, se tiene

$$\begin{aligned} 2\alpha + \gamma &= 0, \\ 2\alpha + \beta &= 0, \end{aligned}$$

y si $x = \pi/2$

$$\begin{aligned} 2\alpha + \beta &= 0, \\ -\gamma &= 0; \end{aligned}$$

de donde $\alpha = \beta = \gamma = 0$. Luego f_1, f_2 y f_3 son L.I. en $\mathscr{F}(X, \mathbb{C})$. ✔

3.6.2 Ejercicios propuestos

El lector encontrará la respuesta a los ejercicios en cursiva en el apéndice E al final del libro.

Geometría de los espacios $\mathbb{R}^n$ (respuestas en página 1077)

En los ejercicios 1 a 21, $\vec{u} = (-1, 2, 4)$, $\vec{v} = (-1, 0, 2)$ y $\vec{w} = (-3, 4, 5)$.

1 Encontrar $-\vec{u}$.

2 Calcular $\|\vec{w}\|$.

3 Hallar $\vec{u} + \vec{v}$.

4 Encontrar $\vec{u} - 3\vec{v} + 7\vec{w}$.

5 Calcular $\vec{v} - 5\vec{w}$.

6 Hallar $\frac{2}{5}\vec{v}$.

7 Calcular $\vec{w} - \vec{u}$.

8 Encontrar un vector paralelo a $\vec{u}$ con la misma dirección.

9 Hallar un vector paralelo a $\vec{v}$ con dirección opuesta.

10 Hallar $\vec{w} \cdot \vec{v}$

11 Calcular $\vec{u} \cdot (3\vec{v} + \vec{w})$.

12 Encontrar $(\vec{u} + \vec{v}) \cdot \vec{w}$.

13 Encontrar el ángulo entre $\vec{u}$ y $\vec{v}$.

14 Calcular el ángulo entre $\vec{v}$ y $\vec{w}$.

15 Encontrar el valor de x para que $(x,-1,4)$ sea ortogonal a $\vec{u}$.

16 Encontrar el valor de y para que $(2,y,-3)$ sea ortogonal a $\vec{w}$.

17 Encontrar un vector no nulo que sea ortogonal a $\vec{u}$ y a $\vec{v}$ utilizando el concepto de producto punto.

18 Encontrar un vector no nulo que sea ortogonal a $\vec{v}$ y a $\vec{w}$ utilizando el concepto de producto vectorial dado en los ejercicios resueltos 10 y 11 de este capítulo.

19 Calcular $\vec{u}\times\vec{v}$ y $\vec{v}\times\vec{u}$. (Utilizar la definición de producto vectorial dada en el ejercicio resuelto 11 de este capítulo.)

20 Encontrar $(\vec{w}\times\vec{v})\cdot\vec{u}$.

21 Hallar $(\vec{u}\times\vec{w})\times\vec{v}$.

En los ejercicios 22 a 25, $\vec{u}\times\vec{v}$ es el producto vectorial (cruz) definido en los ejercicios resueltos 10 y 11, $\vec{u}$, $\vec{v}$, $\vec{w}$ son vectores cualesquiera de $\mathbb{R}^3$, y α es un número real arbitrario.

22 Demostrar que $\vec{u}\times\vec{v}=-(\vec{v}\times\vec{u})$.

23 Probar que $\vec{u}\times(\vec{v}+\vec{w})=\vec{u}\times\vec{v}+\vec{u}\times\vec{w}$.

24 Mostrar que $(\alpha\vec{u})\times\vec{v}=\vec{u}\times(\alpha\vec{v})=\alpha(\vec{u}\times\vec{v})$.

25 Demostrar que $(\vec{u}\times\vec{v})\times\vec{w}=(\vec{u}\cdot\vec{w})\vec{v}-(\vec{v}\cdot\vec{w})\vec{u}$.

En los ejercicios 26 a 39, calcular el área o el volumen de la configuración dada en $\mathbb{R}^2$ o $\mathbb{R}^3$, utilizar la interpretación geométrica del determinante del apartado 3.1.2 (pág. 117) y la definición e interpretación geométrica del producto vectorial dada en los ejercicios resueltos 10, 11 y 13 de este capítulo.

26 El área del paralelogramo generado por los vectores $(2,5)$ y $(6,2)$.

27 El volumen del paralelepípedo generado por los vectores $\vec{u}=(3,-2,4)$, $\vec{v}=(-2,3,5)$ y $\vec{w}=(1,2,5)$.

28 El área del paralelogramo generado por los vectores $(-3,5)$ y $(2,4)$.

29 El volumen del paralelepípedo generado por los vectores $\vec{u}=(-2,2,5)$, $\vec{v}=(3,0,6)$ y $\vec{w}=(2,4,2)$.

30 El área del paralelogramo generado por los vectores $(-1,2,-1)$ y $(2,-3,2)$.

31 El volumen del paralelepípedo con vértices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(1,1,0)$, $(1,1,1)$, $(1,0,1)$, $(0,1,1)$.

32 El volumen del paralelepípedo con vértices $(1,1,4)$, $(2,3,7)$, $(0,3,8)$, $(0,0,7)$, $(1,5,11)$, $(0,4,14)$, $(1,2,10)$, $(-1,2,11)$.

33 El volumen del paralelepípedo con vértices $(2,3,4)$, $(3,5,7)$, $(3,4,1)$, $(5,3,9)$, $(4,6,4)$, $(7,6,9)$, $(6,5,12)$, $(6,4,6)$.

34 El área del paralelogramo con vértices $(2,2)$, $(4,6)$, $(7,3)$ y $(9,7)$.

35 El área del paralelogramo con vértices $(-4,1)$, $(-2,6)$, $(4,3)$ y $(2,-3)$.

36 El área del paralelogramo con vértices $(1,0,1)$, $(3,1,4)$, $(0,2,9)$ y $(-2,1,6)$.

37 El área del triángulo con vértices $(4,0,3)$, $(-1,3,5)$ y $(1,1,-2)$.

38 El área del triángulo con vértices $(1,-1,2)$, $(3,-2,1)$, $(2,3,4)$.

39 El área del triángulo acotado por las rectas $z+x=1$, $y+x=1$ y $y+z=1$.

En los ejercicios 40 a 58, $\vec{u}=(1,-2,-4,3,5)$, $\vec{v}=(-2,3,-1,2,-1)$ y $\vec{w}=(2,-1,-3,1,-2)$.

40 Encontrar $-\vec{w}$.

41 Calcular $\|\vec{u}\|$.

42 Hallar $\vec{u}+\vec{w}$.

43 Encontrar $2\vec{u}-4\vec{v}+\vec{w}$.

44 Calcular $\vec{v}-3\vec{w}$.

45 Hallar $4\vec{v}$.

46 Calcular $\vec{w}-\vec{u}$.

47 Encontrar un vector paralelo a $\vec{u}$ con la misma dirección.

48 Hallar un vector paralelo a $\vec{v}$ con dirección opuesta.

49 Hallar $\vec{w}\cdot\vec{v}$.

50 Encontrar $\vec{u}\cdot\vec{v}$.

51 Calcular $\vec{u}\cdot(2\vec{v}-3\vec{w})$.

52 Encontrar $(\vec{u}+\vec{v})\cdot\vec{w}$.

53 Encontrar el ángulo entre $\vec{u}$ y $\vec{v}$.

54 Calcular el ángulo entre $\vec{v}$ y $\vec{w}$.

55 Encontrar el valor de x para que $(x,1,4,-1,2)$ sea ortogonal a $\vec{u}$.

56 Encontrar el valor de y para que $(-1,y,3,-2,1)$ sea ortogonal a $\vec{w}$.

57 Encontrar un vector no nulo que sea ortogonal a $\vec{u}$ y a $\vec{v}$.

58 Hallar un vector no nulo que sea ortogonal a $\vec{u}$ y a $\vec{w}$.

59 Si $\vec{u}=(-1,2,3)$ y $\vec{v}=(2,-1,-1)$, hallar $\mathrm{d}(\vec{u},\vec{v})$ en $\mathbb{R}^3$.

60 Si $\vec{u}=(2,-1,1,2)$ y $\vec{v}=(2,-3,4,-2)$, hallar $\mathrm{d}(\vec{u},\vec{v})$ en $\mathbb{R}^4$.

61 Encontrar la distancia entre los vectores $(-3,-1,2,-2,1)$ y $(5,-2,4,0,1)$ de $\mathbb{R}^4$.

62 Hallar la distancia entre los vectores $(-1,0,-2,5,-6,3,2)$ y $(4,1,2,6,2,-2,1)$ de $\mathbb{R}^7$.

63 Demostrar la propiedad de distributividad del producto punto en $\mathbb{R}^n$ (propiedad 3, pág. 123); esto es,

$$\vec{u}\cdot(\vec{v}+\vec{w})=\vec{u}\cdot\vec{v}+\vec{u}\cdot\vec{w} \quad \forall\vec{u},\vec{v},\vec{w}\in\mathbb{R}^n.$$

64 Demostrar la propiedad de positividad del producto punto en $\mathbb{R}^n$ (propiedad 4, pág. 123); esto es,

$$\begin{aligned}\vec{u}\cdot\vec{u} &\geq 0 \quad \forall\vec{u}\in\mathbb{R}^n \text{ y}\\ \vec{u}\cdot\vec{u} &= 0 \Leftrightarrow \vec{u}=\vec{0}_{\mathbb{R}^n}.\end{aligned}$$

65 Considerar la desigualdad de Schwarz (cfr. teorema 3.1, pág. 124) para el caso $\mathbb{R}^3$, con $\vec{u}=(x_1,x_2,x_3)$ y $\vec{v}=(y_1,y_2,y_3)$. En el caso de que uno de estos vectores sea el neutro aditivo, la desigualdad de Schwarz se traduce a igualdad. Si ninguno de estos vectores es nulo, probar que hay igualdad en la desigualdad (3.6) si y sólo si uno de los vectores es múltiplo escalar del otro, por ejemplo $\vec{v}=\lambda\vec{u}$. Utilizar la definición algebraica de producto punto prescindiendo del concepto de ángulo entre vectores.

66 Demostrar que, en general, para cualquier par de vectores no nulos en $\mathbb{R}^n$ $|\vec{u}\cdot\vec{v}|=\|\vec{u}\|\,\|\vec{v}\|$ si y sólo si uno de ellos es múltiplo escalar del otro, por ejemplo $\vec{v}=\lambda\vec{u}$.

67 Probar las propiedades 1, 2 y 3 del valor absoluto enunciadas en la página 126.

68 Probar que para todo par de números reales, x y y, se cumple:

(a) $||x|-|y||\leq|x-y|$.

(b) $|x-y|\leq|x|+|y|$.

69 Encontrar la ecuación del plano que pasa por el punto $\vec{u}=(-1,2,5)$ y es ortogonal al vector $\vec{N}=(-3,2,4)$.

70 Hallar la ecuación del plano que pasa por el punto $\vec{u} = (2, -1, 4)$ y es ortogonal al vector $\vec{N} = (1, -2, 1)$.

71 Encontrar la ecuación del plano que pasa por el punto $\vec{u} = (3, -2, 4)$ y es ortogonal al vector $\vec{N} = (1, 1, -3)$.

72 Encontrar la ecuación del plano que pasa por el punto $\vec{u} = (2, 3, -4)$ y es ortogonal al vector $\vec{N} = (1, -3, 3)$.

73 Hallar la ecuación del plano que pasa por los puntos $(-2, 1, 3)$, $(1, 2, -2)$ y $(2, -1, 6)$.

74 Encontrar la ecuación del plano que pasa por los puntos $(4, 1, -1)$, $(-2, -1, -1)$ y $(6, -2, 2)$.

75 Encontrar la ecuación del plano que pasa por los puntos $(2, 3, -1)$, $(-1, -2, 1)$ y $(2, -5, 7)$.

76 Hallar la ecuación del plano que pasa por los puntos $(3, 1, 2)$, $(2, 3, 7)$ y $(1, -3, 8)$.

Espacios vectoriales (respuestas en páginas 1077-1078)

En los ejercicios 77 a 89, considere los conjuntos dados con las operaciones usuales de suma entre sus elementos y multiplicación por un escalar. Determine si *(a)* la suma es cerrada y *(b)* la multiplicación es cerrada en cada conjunto.

77 La recta $y = 2x - 1$.

78 El primer cuadrante del plano cartesiano $\mathbb{R}^2$.

79 El conjunto de matrices cuadradas de orden 3 triangulares superiormente.

80 El conjunto de matrices simétricas de orden n.

81 El conjunto de funciones $f : \mathbb{R} \to \mathbb{R}$ que son periódicas con periodo T $(f(x+T) = f(x)\ \forall x)$.

82 El conjunto de matrices invertibles de orden 3.

83 El conjunto de vectores en $\mathbb{R}^3$ que están contenidos en el plano x, z.

84 El conjunto de puntos en el plano $\mathbb{R}^2$ cuyas componentes son números racionales.

85 El rectángulo $Q = [a, b] \times [c, d]$; es decir, el conjunto de puntos (x, y) tales que $a \le x \le b$ y $c \le y \le d$, donde $a < b$ y $c < d$.

86 Los puntos dentro y sobre la esfera $x^2 + y^2 + z^2 = 1$; esto es, el conjunto de puntos (x, y, z) cuya distancia al origen es menor o igual a 1.

87 El conjunto de matrices 2×2 que tienen todas sus componentes nulas excepto una.

88 El conjunto de matrices 3×3 cuya traza es cero; es decir, $\sum_{i=1}^{3} a_{ii} = 0$.

89 El conjunto de números reales que tienen la forma $a+b\sqrt{2}$, con $a,b\in\mathbb{R}$.

En los ejercicios 90 a 97, determinar los conjuntos que, junto con las operaciones suma y multiplicación por un escalar ahí indicadas, son espacios vectoriales. Para los que resulten ser espacios vectoriales probar los 10 axiomas correspondientes, y para los que no, indicar mediante contraejemplos los axiomas que no se cumplen.

90 $\mathbb{R}^3$ con la suma usual de vectores, pero con la multiplicación por un escalar r definida por $r(x,y,z) = (rz,ry,rx)$.

91 $\mathbb{R}^2$ con el producto usual de un escalar por un vector, pero con la suma definida por $(x,y)\dot{+}(x_1,y_1) = (y+y_1, x+x_1)$.

92 El conjunto de matrices de orden 3 con la suma usual de matrices pero con multiplicación por un escalar definida, para cualquier matriz A, por $rA = \mathscr{O}$, donde $\mathscr{O}$ es la matriz cero 3×3.

93 $\mathscr{F}(X)$ con el producto usual de un escalar por una función pero con suma definida por $(f\dot{+}g)(x) = \text{máx}\{f(x),g(x)\}$.

94 El conjunto de matrices $n\times n$ triangulares superiormente con las operaciones usuales.

95 El conjunto de matrices cuadradas de orden n que son diagonales con las operaciones usuales.

96 $\mathbb{R}^2$ con la suma usual de vectores pero con el producto de un escalar por un vector definido por

$$\alpha(x,y) = \begin{cases} (0,0) & \text{si } \alpha = 0 \\ (\alpha x, \frac{1}{\alpha}y) & \text{si } \alpha \neq 0 \end{cases}$$

97 $\mathbb{R}^2$ con la multiplicación usual de un escalar por un vector pero con la suma definida por

$$(x,y)\dot{+}(x_1,y_1) = (x+2x_1, y+3y_1).$$

98 Demostrar que $\mathbb{R}^\infty$, el conjunto de sucesiones (cfr. ejemplo 3.12, pág. 133), junto con las operaciones usuales de suma de sucesiones y multiplicación de un escalar por una sucesión, es un espacio vectorial.

99 Mostrar que **P**, el conjunto de polinomios, con las operaciones usuales de suma de polinomios y multiplicación de un escalar por un polinomio es un espacio vectorial (cfr. ejemplo 3.11).

100 Probar la propiedad 9 del teorema 3.4 (cfr. pág. 138).

101 Si **E** es un espacio vectorial, probar que

(a) $(-\alpha)\vec{u} = -(\alpha\vec{u}) = \alpha(-\vec{u})$, para todo $\vec{u}\in\mathbf{E}$ y para todo escalar α.

(b) $\alpha(\vec{u}-\vec{v}) = \alpha\vec{u}-\alpha\vec{v}$, para todo par $\vec{u},\vec{v}\in\mathbf{E}$ y para todo $\alpha\in\mathbb{R}$.

(c) Si $\alpha\vec{u} = \alpha\vec{v}$ y $\alpha \neq 0$, entonces $\vec{u} = \vec{v}$.

(d) Si $\alpha\vec{u} = \beta\vec{u}$ y $\vec{u} \neq \vec{0}_{\mathbf{E}}$, entonces $\alpha = \beta$.

(e) $-(\vec{u}+\vec{v}) = (-\vec{u})+(-\vec{v})$ para todo par de vectores $\vec{u},\vec{v} \in \mathbf{E}$.

(f) $(\alpha+\beta)(\vec{u}+\vec{v}) = \alpha\vec{u}+\alpha\vec{v}+\beta\vec{u}+\beta\vec{v}$, para todo par de vectores $\vec{u},\vec{v} \in \mathbf{E}$ y para todo par de escalares α, β.

102 Sea $X = \{0,1\}$ y $\mathscr{F}(X)$ el espacio de funciones $f : X \to \mathbb{R}$. Si $f(x) = 3x+1$, $g(x) = 1+4x-x^2$ y $h(x) = x6^x+2$, probar que:

(a) $f = g$.

(b) $f+g = h$.

En los ejercicios 103 a 116, determinar si el subconjunto S es un subespacio del espacio dado. En caso de una respuesta afirmativa, demostrar rigurosamente la afirmación; y en caso de una respuesta negativa, indicar con un contraejemplo la propiedad o propiedades de subespacio vectorial que no se cumplen.

103 $S = \{(-x,y)\,|\,x,y \in \mathbb{R}\}$ en $\mathbb{R}^2$.

104 $S = \{\vec{u} \in \mathbb{R}^n\,|\,\text{las componentes de } \vec{u} \text{ son números racionales}\}$ en $\mathbb{R}^2$.

105 $S = \{(x,y)\,|\,xy \geq 0\}$ en $\mathbb{R}^2$.

106 $S = \{(x,y,z)\,|\,2x+3y+z = 4\}$ en $\mathbb{R}^3$.

107 $S = \{(x_1,x_2,x_3,x_4)\,|\,x_1-3x_2+2x_3-x_4 = 0\}$ en $\mathbb{R}^4$.

108 $S = \{(x,y,z)\,|\,x^2+y^2+z^2 \leq 16\}$ en $\mathbb{R}^3$.

109 $S = \{(x,y)\,|\,-1 \leq x \leq 1, -1 \leq y \leq 1\}$ en $\mathbb{R}^2$.

110 S, el conjunto de matrices triangulares superiores, en el espacio $\mathfrak{M}_n$.

111 S, el conjunto de matrices invertibles, en el espacio $\mathfrak{M}_n$.

112 $S = \{(a_n)\,|\,\text{lím}_{n\to\infty} a_n \text{ existe}\}$ en $\mathbb{R}^\infty$ (el espacio de sucesiones).

113 $S = \{(a_n)\,|\,\text{lím}_{n\to\infty} a_n = 0\}$ en $\mathbb{R}^\infty$.

114 $S = \{(a_n)\,|\,\text{lím}_{n\to\infty} a_n = 1\}$ en $\mathbb{R}^\infty$.

115 $S = \{(a_n)\,|\,a_n+3a_{n-1}+2a_{n-2} = 0,\ n \geq 2\}$ en $\mathbb{R}^\infty$.

116 $S = \{A\,|\,\text{tales que tra}(A) = 0\}$ en $\mathfrak{M}_n$, donde $\text{tra}(A) = \sum_{i=1}^{n} a_{ii}\,\mathfrak{M}$.

En los ejercicios 117 a 129, demostrar que el conjunto dado es un subespacio del espacio vectorial indicado. La continuidad y derivabilidad de las funciones en los extremos del intervalo se considera en forma lateral; es decir, continuidad y derivabilidad por la derecha en a y continuidad y derivabilidad por la izquierda en b.

117 $C^1[a,b] = \{f : [a,b] \to \mathbb{R} \,|\, f$ es derivable con continuidad en $[a,b]\}$ en $C[a,b]$, el espacio de funciones continuas en $[a,b]$.

118 $C^n[a,b] = \{f : [a,b] \to \mathbb{R} \,|\, f$ tiene derivada hasta el orden n en $[a,b]\}$

119 $C^\infty[a,b] = \{f : [a,b] \to \mathbb{R} \,|\, f$ tiene derivada de todo orden en $[a,b]\}$ en $C[a,b]$.

120 $S = \{f \,|\, a_2 f'' + a_2 f' + a_1 f = \theta\}$, donde θ es la función constante cero, en $C^1[a,b]$.

121 $S = \{(a_n) \,|\, \sum_{n=1}^{\infty} a_n$ converge$\}$ en $\mathbb{R}^\infty$.

122 $S = \{(a_n) \,|\, \sum_{n=1}^{\infty} |a_n|$ converge$\}$ en $\mathbb{R}^\infty$.

123 $S = \{f : [a,b] \to \mathbb{R} \,|\, \int_a^b f(x)dx = 0\}$ en $C[a,b]$.

124 $S = \{f : \mathbb{R} \to \mathbb{R} \,|\, f$ es par: $f(-x) = f(x)\,\forall x\}$ en $\mathscr{F}(\mathbb{R})$.

125 $S = \{f : \mathbb{R} \to \mathbb{R} \,|\, f$ es impar: $f(-x) = -f(x)\,\forall x\}$ en $\mathscr{F}(\mathbb{R})$.

126 $S = \{p \,|\, p$ es un polinomio y las potencias pares de x son nulas$\}$ en $\mathbf{P}$.

127 $S = \{A \,|\, A$ es una matriz cuadrada de orden n triangular inferior$\}$.

128 $S = \{f : [0,1] \to \mathbb{R} \,|\, f(0) = f'(0) = 0\}$ en $C^1[0,1]$.

129 $S = \{f : X \to \mathbb{R} \,|\, f$ alcanza sólo un número finito de valores en $X\}$ (es decir, la imagen de cualquier función $f \in S$ es un conjunto finito; o sea que el conjunto de los $f(x)$, con $x \in X$, es finito para cada $f \in S$) en $\mathscr{F}(X)$.

130 Sean

$$\begin{aligned} S_1 &= \{(a,b,c) \in \mathbb{R}^3 \,|\, a = c, b = 3c\}. \\ S_2 &= \{(a,b,c) \in \mathbb{R}^3 \,|\, a + b + c = 0\}. \\ S_3 &= \{(a,b,c) \in \mathbb{R}^3 \,|\, a - 2b + 3c = 0\}. \end{aligned}$$

(a) Probar que S_1, S_2 y S_3 son subespacios de $\mathbb{R}^3$.

(b) Por el ejercicio resuelto 30, $S_i \cap S_j$ es un subespacio de $\mathbb{R}^3$ para cualquier combinación de los subíndices i, j.

(i) Describir el subespacio $S_1 \cap S_2$.

(ii) Describir el subespacio $S_1 \cap S_3$.

(iii) Describir el subespacio $S_2 \cap S_3$.

131 Probar que los únicos subespacios propios (distintos de $\{\vec{0}_{\mathbb{R}^2}\}$) de $\mathbb{R}^2$ son las rectas que pasan por el origen.

132 Probar que los únicos subespacios propios (distintos de $\{\vec{0}_{\mathbb{R}^3}\}$) de $\mathbb{R}^3$ son las rectas y los planos que pasan por el origen.

133 (**Producto de espacios vectoriales**). Sean $\mathbf{E}$ y $\mathbf{F}$ un par de espacios vectoriales reales. Se definen

$$\mathbf{E}\times\mathbf{F}=\{(\vec{u},\vec{v})\,|\,\vec{u}\in\mathbf{E}\text{ y }\vec{v}\in\mathbf{F}\}$$

y, para todo $(\vec{u},\vec{v}),(\vec{u}_1,\vec{v}_1)\in\mathbf{E}\times\mathbf{F}$, $\alpha\in\mathbb{R}$,

$$\begin{aligned}(\vec{u},\vec{v})+(\vec{u}_1,\vec{v}_1) &= (\vec{u}+\vec{u}_1,\vec{v}+\vec{v}_1),\\ \alpha(\vec{u},\vec{v}) &= (\alpha\vec{u},\alpha\vec{v}).\end{aligned}$$

Donde $\vec{u}+\vec{u}_1$, $\alpha\vec{u}$ y $\vec{v}+\vec{v}_1$, $\alpha\vec{v}$ son las operaciones suma de vectores y multiplicación por un escalar en $\mathbf{E}$ y $\mathbf{F}$, respectivamente. Probar que $\mathbf{E}\times\mathbf{F}$ es un espacio vectorial con estas operaciones.

134 Probar que $\mathbb{R}$ es un espacio vectorial con las operaciones usuales de suma de números reales y multiplicación de números reales.

135 ¿Qué espacio vectorial es $\mathbb{R}\times\mathbb{R}$?

136 ¿Qué espacio vectorial es $\underbrace{\mathbb{R}\times\mathbb{R}\times\cdots\times\mathbb{R}}_{n}$?

137 (**Complemento ortogonal**). Sea $S<\mathbb{R}^n$. Probar que

$$S^{\perp}=\{\vec{u}\in\mathbb{R}^n\,|\,\vec{u}\perp\vec{v}\text{ para todo }\vec{v}\in S\}$$

es un subespacio de $\mathbb{R}^n$, el llamado complemento ortogonal de S.

138 Sean $\vec{u}=(1,1,1)$ y $S=\text{gn}(\vec{u})$, determinar $S^{\perp}$, el complemento ortogonal de S definido en el ejercicio precedente.

139 Probar que si S_1,S_2 son subespacios de un espacio vectorial $\mathbf{E}$ y $S_1\cup S_2$ es también un subespacio de $\mathbf{E}$, entonces $S_1\subset S_2$ o $S_2\subset S_1$.

140 Sean $\mathbf{E}$ un espacio vectorial y S_1,S_2 un par de subespacios. Mostrar que si $\mathbf{E}=S_1\cup S_2$, entonces $S_1=\mathbf{E}$ o $S_2=\mathbf{E}$.

141 Sean S_1 y S_2 los conjuntos de matrices cuadradas de orden n que son triangulares superior e inferior, respectivamente. Mostrar que S_1 y S_2 son subespacios de $\mathfrak{M}_n$ y que $\mathfrak{M}_n=S_1\oplus S_2$, la suma directa de S_1 y S_2 (cfr. ejercicio resuelto 27).

142 Sean $S_1=\{(a_1,\ldots,a_{n-1},0)\,|\,a_i\in\mathbb{R},\,i=1,\ldots,n-1\}$ y $S_2=\{(a_1,\ldots,a_{n-1},a_n)\,|\,a_i=0,\,i=1,\ldots,n-1\}$. Probar que S_1,S_2 son subespacios de $\mathbb{R}^n$ y que $\mathbb{R}^n=S_1\oplus S_2$ (cfr. ejercicio resuelto 27).

143 Sean $S_1 = \{f : \mathbb{R} \to \mathbb{R} \,|\, f$ es par: $f(-x) = f(x)\ \forall x \in \mathbb{R}\}$ y $S_2 = \{f : \mathbb{R} \to \mathbb{R} \,|\, f$ es impar: $f(-x) = -f(x)\ \forall x \in \mathbb{R}\}$. Probar que S_1 y S_2 son subespacios de $\mathscr{F}(\mathbb{R})$, el espacio de funciones $f : \mathbb{R} \to \mathbb{R}$, y que $\mathscr{F}(\mathbb{R}) = S_1 \oplus S_2$.

En los ejercicios 144 a 159 escribir, si es posible, el vector $\vec{v}$ como combinación lineal de los vectores $\vec{u}_i$.

144 $\vec{v} = (-3, 2)$; $\vec{u}_1 = (1, -1)$ y $\vec{u}_2 = (3, 2)$ en $\mathbb{R}^2$.

145 $\vec{v} = (1, 5)$; $\vec{u}_1 = (-1, 2)$, $\vec{u}_2 = (-3, 6)$ en $\mathbb{R}^2$.

146 $\vec{v} = (-1, 1, 4)$; $\vec{u}_1 = (1, -1, 3)$, $\vec{u}_2 = (3, -2, 2)$ y $\vec{u}_3 = (2, -1, 3)$ en $\mathbb{R}^3$.

147 $\vec{v} = (-11, 6, 7, -2)$; $\vec{u}_1 = (1, 0, -2, 4)$ y $\vec{u}_2 = (-3, 2, 1, 2)$ en $\mathbb{R}^4$.

148 $\vec{v} = (8, -10, -1, -9, -17)$, $\vec{u}_1 = (1, 1, 1, 0, -1)$, $\vec{u}_2 = (-2, 3, 1, 2, 3)$, $\vec{u}_3 = (-7, 8, 2, 6, 10)$, $\vec{u}_4 = (-2, 0, 1, -1, -3)$ en $\mathbb{R}^5$.

149 $\vec{v} = 4$; $\vec{u}_1 = \cos^2 x$ y $\vec{u}_2 = \operatorname{sen}^2 x$ en $\mathscr{F}(\mathbb{R})$, el espacio de funciones $f : \mathbb{R} \to \mathbb{R}$.

150 $\vec{v} = \cos 4x$; $\vec{u}_1 = \cos^2 2x$, $\vec{u}_2 = \operatorname{sen}^2 2x$ en $\mathscr{F}(\mathbb{R})$.

151 $\vec{v} = \cos 2x$; $\vec{u}_1 = 1$, $\vec{u}_2 = \cos^2 x$ en $\mathscr{F}(\mathbb{R})$.

152 $\vec{v} = e^{-x}$; $\vec{u}_1 = \cosh x$, $\vec{u}_2 = e^x$ en $\mathscr{F}(\mathbb{R})$.

153 $\vec{v} = \operatorname{senh} x$; $\vec{u}_1 = e^x$, $\vec{u}_2 = e^{-x}$ en $\mathscr{F}(\mathbb{R})$.

154 $\vec{v} = 3x^2 - 2x + 1$; $\vec{u}_1 = 2$, $\vec{u}_2 = 1 - x$, $\vec{u}_3 = (x-2)^2$ en $\mathbf{P}$.

155 $\vec{v} = \begin{bmatrix} -12 & 4 \\ 2 & -1 \end{bmatrix}$; $\vec{u}_1 = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$, $\vec{u}_2 = \begin{bmatrix} -1 & 0 \\ 2 & 0 \end{bmatrix}$, $\vec{u}_3 = \begin{bmatrix} -3 & 1 \\ 0 & 0 \end{bmatrix}$ en $\mathfrak{M}_2$.

156 $\vec{v} = \dfrac{1 - x^{n+1}}{1 - x}$, $\vec{u}_i = x^i$, $i = 0, 1, \ldots, n$ en $\mathscr{F}(X)$, donde $X = \mathbb{R} - \{1\}$.

157 $\vec{v} = \dfrac{\operatorname{sen}(\frac{7}{2}x)}{2\operatorname{sen}(\frac{1}{2}x)}$; $\vec{u}_1 = 1$, $\vec{u}_2 = \cos x$, $\vec{u}_3 = \cos 2x$, $\vec{u}_4 = \cos(3x)$ en $\mathscr{F}((0, 2\pi))$.

158 $\vec{v} = \operatorname{sen} 3x \cos x$; $\vec{u}_1 = \operatorname{sen} 4x$, $\vec{u}_2 = \operatorname{sen} 2x$ en $\mathscr{F}(\mathbb{R})$.

159 $\vec{v} = \cos 5x \cos x$; $\vec{u}_1 = \cos 6x$, $\vec{u}_2 = \cos 4x$ en $\mathscr{F}(\mathbb{R})$.

En los ejercicios 160 a 162, $\vec{u}_1 = (-3, 1, 2)$, $\vec{u}_2 = (2, -1, 3)$ y $\vec{u}_3 = (-1, 1, 2)$.

160 Determinar si $\vec{v} = (14, -6, -2) \in \operatorname{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$. En caso afirmativo escribir $\vec{v}$ como combinación lineal de los vectores $\vec{u}_i$.

161 Determinar si $\vec{v} = (-8, 4, -2) \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$. En caso afirmativo escribir $\vec{v}$ como combinación lineal de los vectores $\vec{u}_i$.

162 Determinar si $\text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3) = \mathbb{R}^3$.

En los ejercicios 163 a 165, $\vec{u}_1 = (-1, 1, 3)$, $\vec{u}_2 = (3, 1, -2)$ y $\vec{u}_3 = (-5, 1, 8)$.

163 Determinar si $\vec{v} = (-10, 2, 16) \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$. En caso afirmativo escribir $\vec{v}$ como combinación lineal de los vectores $\vec{u}_i$.

164 Determinar si $\vec{v} = (-2, -2, -1) \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$. En caso afirmativo escribir $\vec{v}$ como combinación lineal de los vectores $\vec{u}_i$.

165 Determinar si $\text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3) = \mathbb{R}^3$.

En los ejercicios 166 a 168, $A_1 = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, A_3 = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, A_4 = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$.

166 Determinar si

$$\vec{v} = \begin{bmatrix} -6 & 1 \\ 0 & -1 \end{bmatrix} \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4).$$

En caso afirmativo escribir $\vec{v}$ como combinación lineal de los vectores $\vec{u}_i$.

167 Determinar si

$$\vec{v} = \begin{bmatrix} -4 & 1 \\ 1 & 0 \end{bmatrix} \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4).$$

En caso afirmativo escribir $\vec{v}$ como combinación lineal de los vectores $\vec{u}_i$.

168 Determinar si $\text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4) = \mathfrak{M}_2$.

En los ejercicios 169 a 171, $A_1 = \begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix}, A_3 = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}, A_4 = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix}$.

169 Determinar si

$$\vec{v} = \begin{bmatrix} -4 & -2 \\ -8 & 2 \end{bmatrix} \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4).$$

En caso afirmativo escribir $\vec{v}$ como combinación lineal de los vectores $\vec{u}_i$.

170 Determinar si

$$\vec{v} = \begin{bmatrix} 0 & 2 \\ 4 & -2 \end{bmatrix} \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4).$$

En caso afirmativo escribir $\vec{v}$ como combinación lineal de los vectores $\vec{u}_i$.

171 Determinar si $\text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4) = \mathfrak{M}_2$.

En los ejercicios 172 a 180, determinar si los vectores indicados generan al espacio dado.

172 $(-1,0,2), (-1,1,1); \mathbb{R}^3$.

173 $(-1,1), (3,-3), (4,-2); \mathbb{R}^2$.

174 $(2,1), (1,1); \mathbb{R}^2$.

175 $(1,-1,0), (-1,2,3), (-1,1,2); \mathbb{R}^3$.

176 $(2,-1,3), (0,1,1), (2,0,3); \mathbb{R}^3$.

177 $(1,1,0), (2,-1,2), (3,-1,1); \mathbb{R}^3$.

178 $(2,1,0,-1), (2,-1,3,0), (-2,3,1,4); \mathbb{R}^4$.

179 $(2,1,0,-1), (2,-1,3,0), (-2,3,1,4), (0,-1,1,1); \mathbb{R}^4$.

180 $(1,0,-1,1), (2,1,-1,1), (3,-2,4,2), (0,0,-1,1); \mathbb{R}^4$.

En los ejercicios 181 a 188, determinar si los vectores generan al espacio dado; de ser así, reducir el conjunto de vectores a un mínimo de generadores.

181 $(-1,2), (-1,1), (3,2), (-1,-1); \mathbb{R}^2$.

182 $(1,-2,0), (-1,1,2), (1,-4,4), (1,2,-8); \mathbb{R}^3$.

183 $(4,-1,2), (-2,1,5), (-1,3,1), (1,0,2); \mathbb{R}^3$.

184 $(2,-3,2), (1,-2,1), (-4,5,-4), (-4,7,-4); \mathbb{R}^3$.

185 $(1,0,1), (-1,1,0), (2,-1,3), (2,-1,1), (-1,2,2); \mathbb{R}^3$.

186 $(1,-2,1,2), (-1,0,2,3), (-2,1,3,2), (-1,1,1,0), (-2,1,2,1), (-1,2,1,3); \mathbb{R}^4$.

187 $(1,0,-1,1), (2,1,0,-1), (-1,1,3,2), (3,0,-3,2), (2,1,-2,4,); \mathbb{R}^4$.

188 $(-2,3,1,2), (1,-2,2,3), (2,3,0,0), (-3,4,4,7), (-4,6,2,4); \mathbb{R}^4$.

189 Sean $S_1 = \text{gn}((2,1,1),(1,3,2))$ y $S_2 = \text{gn}((1,-1,0),(-3,1,0))$. Por el ejercicio resuelto 30, $S_1 \cap S_2 < \mathbb{R}^3$. Encontrar un conjunto generador del subespacio $S_1 \cap S_2$.

En los ejercicios 190 a 195, describir el subespacio generado por los conjuntos vectores dados.

190 $\{x, 1-x, x^2\}$ en **P**, el espacio de polinomios.

191 $\left\{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}$ en $\mathfrak{M}_2$.

192 $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ en $\mathfrak{M}_2$.

193 $\{\vec{e}_i = (a_j) \,|\, a_j = 1$ si $i = j$ y $a_j = 0$ si $i \neq j\}$ en $\mathbb{R}^\infty$, el espacio de sucesiones reales.

194 $\{(k+x)^k \,|\, k = 0, 1, 2, \ldots\}$ en el espacio de polinomios **P**.

195 $\{1, x, x^2, \ldots, x^n, (x+1)^n, (x+2)^n, \ldots, (x+k)^n, \ldots\}$ en el espacio de polinomios **P**.

196 Probar que $\text{gn}(1, x) = \text{gn}(1-3x, x)$ en el espacio de polinomios **P**.

197 Mostrar que si $\vec{u}, \vec{v}$ son cualquier par de vectores en un espacio vectorial **E**, entonces:

(*a*) $\text{gn}(\vec{u}, \vec{v}) = \text{gn}(3\vec{u} - \vec{v}, \vec{v})$.

(*b*) $\text{gn}(\vec{u}, \vec{v}) = \text{gn}(\vec{u} - \vec{v}, \vec{v} + \vec{u},)$.

198 Sean $\vec{u}_i$, $i = 1, \ldots, k$ y $\vec{v}_j$, $j = 1, \ldots, m$, vectores en un espacio vectorial **E**. Establecer condiciones necesarias y suficientes, en términos del concepto de combinación lineal, para que $\text{gn}(\vec{u}_1, \ldots, \vec{u}_k) = \text{gn}(\vec{v}_1, \ldots, \vec{v}_m)$.

199 Sea S un subconjunto de un espacio vectorial **E**. Probar que S es un subespacio de **E** si y sólo si $\mathscr{L}(S)$, el espacio generado por el conjunto S (cfr. ejercicio resuelto 31 de este capítulo), es igual a S.

200 Sea **E** un espacio vectorial y $S_1, S_2 \subset$ **E**. Probar que $S_1 \subset S_2 \Rightarrow \mathscr{L}(S_1) \subset \mathscr{L}(S_2)$ (cfr. ejercicio resuelto 31 de este capítulo). Mostrar que si además $\mathscr{L}(S_1) =$ **E**, entonces también $\mathscr{L}(S_2) =$ **E**.

201 Probar que si S_1, S_2 son cualquier par de subconjuntos de un espacio **E**, entonces $\mathscr{L}(S_1 \cup S_2) = \mathscr{L}(S_1) + \mathscr{L}(S_2)$ (cfr. ejercicio resuelto 31 de este capítulo).

202 Demostrar que $\mathscr{L}(S_1 \cap S_2) \subset \mathscr{L}(S_1) \cap \mathscr{L}(S_2)$ para cualquier par de subconjuntos S_1, S_2 de un espacio vectorial **E** (cfr. ejercicio resuelto 31 de este capítulo).

Dependencia e independencia lineales (respuestas en página 1078)

203 Dar una condición geométrica para que dos vectores no nulos sean L.D. en $\mathbb{R}^2$.

204 Dar una condición geométrica para que dos vectores no nulos sean L.I. en $\mathbb{R}^2$.

205 Argumentar geométricamente por qué tres vectores no nulos son L.D. en $\mathbb{R}^2$.

206 Dar una condición geométrica para que dos vectores no nulos sean L.D. en $\mathbb{R}^3$.

207 Dar una condición geométrica para que dos vectores no nulos sean L.I. en $\mathbb{R}^3$.

208 Dar una condición geométrica para que tres vectores no nulos sean L.D. en $\mathbb{R}^3$.

209 Dar una condición geométrica para que tres vectores no nulos sean L.I. en $\mathbb{R}^3$.

210 Argumentar geométricamente por qué cuatro vectores no nulos son L.D. en $\mathbb{R}^3$.

En los problemas 211 a 257, determinar si el conjunto dado de vectores es L.D. o L.I.

211 $\{(2,8),(-1,-4)\}$ en $\mathbb{R}^2$.

212 $\{(1,-2),(2,-1)\}$ en $\mathbb{R}^2$.

213 $\{(-6,9),(2,-3)\}$ en $\mathbb{R}^2$.

214 $\{(-1,3),(2,3)\}$ en $\mathbb{R}^2$.

215 $\{(-1,1),(2,4),(2,-3)\}$ en $\mathbb{R}^2$.

216 $\{(1,-1,2),(-2,2,-4)\}$ en $\mathbb{R}^3$.

217 $\{(2,-1,3),(4,-2,8)\}$ en $\mathbb{R}^3$.

218 $\{(1,-1,1),(2,-1,3),(2,2,1)\}$ en $\mathbb{R}^3$.

219 $\{(2,4,3),(-5,-2,-5),(4,-8,1)\}$ en $\mathbb{R}^3$.

220 $\{(3,0,-1),(1,2,4),(-3,2,1)\}$ en $\mathbb{R}^3$.

221 $\{(-2,2,5),(4,-2,1),(-3,2,2)\}$ en $\mathbb{R}^3$.

222 $\{(-3,2,4),(2,-2,5),(-1,1,2),(-2,3,1)\}$ en $\mathbb{R}^3$.

223 $\{(2,-1,3,2),(1,1,2,1)\}$ en $\mathbb{R}^4$.

224 $\{(-1,1,3,2),(2,-2,-6,-4)\}$ en $\mathbb{R}^4$.

225 $\{(2,1,-1,2),(-1,1,1,-1),(2,-2,1,1)\}$ en $\mathbb{R}^4$.

226 $\{(-1,-2,3,2),(-3,-2,1,1),(3,-2,7,4)\}$ en $\mathbb{R}^4$.

227 $\{(1,1,0,1),(-1,0,1,2),(2,-1,1,1)\}$ en $\mathbb{R}^4$.

228 $\{(3,-2,1,1),(-2,1,2,0),(-7,4,3,-1)\}$ en $\mathbb{R}^4$.

229 $\{(-1,1,2,1),(-2,2,1,3),(1,-4,2,1),(-2,1,2,1)\}$ en $\mathbb{R}^4$.

230 $\{(3,-2,1,0),(2,-3,4,1),(1,0,-3,2),(-2,4,-4,-4)\}$ en $\mathbb{R}^4$.

231 $\{(2,3,1,2),(1,-2,2,3),(-1,0,2,3),(-2,1,4,2)\}$ en $\mathbb{R}^4$.

232 $\{(3,-4,-5,1),(-1,5,2,6),(-2,3,1,2),(-7,17,10,13)\}$ en $\mathbb{R}^4$.

233 $\{(2,3,1,4),(-1,2,8,7),(-11,3,2,4),(-23,2,1,2),(-2,7,9,13)\}$ en $\mathbb{R}^4$.

234 $\left\{\begin{bmatrix} -1 & 2 & 1 \\ 2 & -3 & 1 \end{bmatrix}, \begin{bmatrix} -3 & 2 & 1 \\ 3 & -2 & 2 \end{bmatrix}, \begin{bmatrix} -1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}\right\}$ en $\mathfrak{M}_{2\times 3}$.

235 $\left\{\begin{bmatrix} 1 & -1 & 2 \\ 3 & -2 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 2 & -4 \\ -5 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 & -2 \\ -1 & -4 & 7 \end{bmatrix}\right\}$ en $\mathfrak{M}_{2\times 3}$.

236 $\left\{\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \begin{bmatrix} 3 & -1 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\right\}$ en $\mathfrak{M}_2$.

237 $\left\{\begin{bmatrix} 3 & -2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} -2 & 0 \\ 0 & -6 \end{bmatrix}\right\}$ en $\mathfrak{M}_2$.

238 $\left\{\begin{bmatrix} 1 & 2 \\ -1 & 2 \\ 1 & 3 \end{bmatrix}, \begin{bmatrix} -1 & 3 \\ 1 & -1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 3 & -1 \\ 2 & 0 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 0 & 3 \\ 2 & 0 \end{bmatrix}\right\}$ en $\mathfrak{M}_{3\times 2}$.

239 $\left\{\begin{bmatrix} -2 & 1 \\ 3 & -2 \\ -1 & -3 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & -1 \\ -3 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ -2 & -3 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 2 & -2 \\ 0 & 3 \end{bmatrix}\right\}$ en $\mathfrak{M}_{3\times 2}$.

240 $\{\cos x, \cos(-x)\}$ en $\mathscr{F}(\mathbb{R})$.

241 $\{\operatorname{sen} x, \operatorname{sen}(-x)\}$ en $\mathscr{F}(\mathbb{R})$.

242 $\{1-x, 2x^2-1, 3x, x^2+2\}$ en $\mathbf{P}$, el espacio de polinomios.

243 $\{3, 2-x, x^3+2, x^2\}$ en $\mathbf{P}$, el espacio de polinomios.

244 $\{1-x, x^2+1, 2+x, x^3, -3+2x-2x^2\}$ en $\mathbf{P}$.

245 $\{1, 1-x, (1-x)^2, (1-x)^3\}$ en $\mathscr{F}(\mathbb{R})$.

246 $\{4, \cos^2 x, \operatorname{sen}^2 x\}$ en $\mathscr{F}(\mathbb{R})$.

247 $\{\cos 4x, \cos^2 2x, \operatorname{sen}^2 2x\}$ en $\mathscr{F}(\mathbb{R})$.

248 $\{\cos 2x, 1, \cos^2 x\}$ en $\mathscr{F}(\mathbb{R})$.

249 $\{e^{-x}, \cosh x, e^x\}$ en $\mathscr{F}(\mathbb{R})$.

250 $\{\operatorname{senh} x, e^x, e^{-x}\}$ en $\mathscr{F}(\mathbb{R})$.

251 $\{(x-1)^{-1}, (x-2)^{-2}, (x-3)^{-3}\}$ en $\mathscr{F}(X)$, donde $X = \mathbb{R} - \{1,2,3\}$.

252 $\{x^{-2}, x^{-1}, 1, x, x^2\}$ en $\mathscr{F}(X)$, donde $X = \mathbb{R} - \{0\}$.

253 $\{(x-1)^{-2}, (x-1)^{-1}, 1, x-1, (x-1)^2\}$ en $\mathscr{F}(X)$, donde $X = \mathbb{R} - \{1\}$.

254 $\{\cos x, \cos 2x, \cos 3x\}$ en $\mathscr{F}(\mathbb{R})$.

255 $\{\operatorname{sen} x, \operatorname{sen} 2x, \operatorname{sen} 3x\}$ en $\mathscr{F}(\mathbb{R})$.

256 $\{\cos x, \operatorname{sen} x, \cos 2x, \operatorname{sen} 2x, \cos 3x, \operatorname{sen} 3x\}$ en $\mathscr{F}(\mathbb{R})$.

257 $\{1, 2+x, 3+2x^2, 4+3x^3\}$ en **P**.

258 Mostrar que si $\vec{u}_1, \vec{u}_2, \vec{u}_3$ son vectores L.I. en un espacio vectorial **E**, entonces los vectores $\vec{v}_1 = 4\vec{u}_1$, $\vec{v}_2 = \vec{u}_1 - 3\vec{u}_2$ y $\vec{v}_3 = 3\vec{u}_1 + 2\vec{u}_3$ son L.I.

259 Probar que si los vectores $\vec{u}_1, \vec{u}_2, \vec{u}_3$ son vectores L.I. en un espacio vectorial **E**, entonces los vectores $\vec{v}_1 = 3\vec{u}_1 + 2\vec{u}_2$, $\vec{v}_2 = \vec{u}_2 - 4\vec{u}_3$ y $\vec{v}_3 = -3\vec{u}_1 - \vec{u}_3$ son L.I.

260 Sean $\vec{u}_1, \vec{u}_2$ y $\vec{u}_3$ vectores de un espacio vectorial **E**.

(a) Probar que los vectores $\vec{u}_1, \vec{u}_2$ son L.I. si y sólo si los vectores $\vec{u}_1 + \vec{u}_2$, $\vec{u}_1 - \vec{u}_2$ son L.I.

(b) Mostrar que los vectores $\vec{u}_1, \vec{u}_2, \vec{u}_3$ son L.I. si y sólo si los vectores $\vec{u}_1 + \vec{u}_2, \vec{u}_1 + \vec{u}_3, \vec{u}_2 + \vec{u}_3$ son L.I.

261 Encontrar escalares α, si existen, tales que los vectores $(1,-2,3)$, $(2,\alpha,-1)$ y $(1,2,1)$ sean L.I.

262 Encontrar escalares α, si existen, tales que los vectores $(1,3,\alpha)$, $(-2,\alpha,1)$ y $(1,2,1)$ sean L.I.

263 En el ejercicio resuelto 39 de este capítulo se probó que si $\vec{u}, \vec{v} \in \mathbb{R}^n$ son L.I. y $A \in \mathfrak{M}_n$ es invertible, entonces $A\vec{u}$ y $A\vec{v}$ son L.I. ¿Es cierta la misma conclusión si la matriz A no es necesariamente invertible?

264 Sea A una matriz cuadrada de orden n y $\vec{u}, \vec{v} \in \mathbb{R}^n$. Probar que si los vectores $A\vec{u}$ y $A\vec{v}$ son L.I., entonces $\vec{u}$ y $\vec{v}$ también son L.I.

Bases y dimensión (respuestas en páginas 1079-1080)

265 Demostrar que el conjunto infinito $\{\cos x, \cos 2x, \cos 3x, \ldots, \cos nx, \ldots\}$ es independiente en el espacio de funciones $\mathscr{F}(J)$, donde J es cualquier intervalo.

266 Probar que el conjunto infinito $\{\operatorname{sen} x, \operatorname{sen} 2x, \operatorname{sen} 3x, \ldots, \operatorname{sen} nx, \ldots\}$ es L.I. en $\mathscr{F}(J)$, donde J es cualquier intervalo.

267 Demostrar que el conjunto infinito

$$\{\cos x, \operatorname{sen} x, \cos 2x, \operatorname{sen} 2x, \cos 3x, \operatorname{sen} 3x, \ldots, \cos nx, \operatorname{sen} nx, \ldots\}$$

es L.I. en $\mathscr{F}(J)$, donde J es cualquier intervalo.

268 Mostrar que el conjunto infinito $\{1, 2+x, 3+2x^2, 4+3x^3, \ldots, n+(n-1)x^{n-1}, \ldots\}$ es L.I. en $\mathscr{F}(\mathbb{R})$.

En los ejercicios 269 a 272, determinar si los conjuntos dados son L.I. o L.D.

269 $\{1, 2+x, 3+2x, \ldots, n+(n-1)x, \ldots\}$ en el espacio de polinomios **P**.

270 $\{1, x, 1+x+x^2, 1+x+x^2+x^3, \ldots, 1+x+x^2+\cdots+x^n, \ldots\}$ en el espacio de polinomios **P**.

271 $\{1, x+2, (x+3)^2, \ldots, (x+n)^{n-1}, \ldots\}$ en el espacio de polinomios **P**.

272 $\{(x-1)^n \,|\, n = 0, \pm 1, \pm 2, \ldots\}$ en el espacio $\mathscr{F}(X)$, donde $X = \mathbb{R} - \{1\}$.

En los ejercicios 273 a 297, determinar si el conjunto de vectores es o no una base del espacio indicado.

273 $\{(-1,1), (2,-2)\}$; $\mathbb{R}^2$.

274 $\{(1,3), (-2,1)\}$; $\mathbb{R}^2$.

275 $\{(1,-2,1), (3,-2,3), (1,1,3)\}$; $\mathbb{R}^3$.

276 $\{(2,-2,1), (3,-1,2), (2,2,2)\}$; $\mathbb{R}^3$.

277 $\{(-4,3,-2), (5,-1,1), (-3,5,-3)\}$; $\mathbb{R}^3$.

278 $\{(2,3,1), (-3,4,1), (-1,1,2)\}$; $\mathbb{R}^3$.

279 $\{(1,2,-3,1), (2,-1,5,2), (-2,3,1,1), (1,0,-1,2)\}$; $\mathbb{R}^4$.

280 $\{(2,-3,2,1), (-1,1,1,0), (3,-2,4,5), (-2,-1,3,-3)\}$; $\mathbb{R}^4$.

281 $\{(-3,0,1,3), (2,-1,1,0), (1,1,1,1), (2,-1,1,7)\}$; $\mathbb{R}^4$.

282 $\{(4,-5,2,1), (2,-2,4,6), (-2,1,3,1), (0,2,-1,4)\}$; $\mathbb{R}^4$.

283 $\{(2,4,-4,6,2), (3,-1,2,1,2), (-1,1,2,5,3), (4,-2,1,1,3), (-1,3,2,4,2)\}$; $\mathbb{R}^5$.

284 $\{(1,-2,-3,1,-2), (-2,1,1,3,2), (-1,1,2,1,1), (-3,0,1,8,1), (-5,4,7,6,5)\}$; $\mathbb{R}^5$.

285 $\left\{\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\right\}$; $\mathfrak{M}_2$.

286 $\left\{\begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}; \mathfrak{M}_2.$

287 $\left\{\begin{bmatrix} 4 & 2 \\ -2 & 4 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} -2 & -6 \\ 0 & -8 \end{bmatrix}, \begin{bmatrix} 1 & -12 \\ -3 & -14 \end{bmatrix}\right\}; \mathfrak{M}_2.$

288 $\left\{\begin{matrix} \begin{bmatrix} -1 & 2 & 2 \\ 3 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & -3 \\ 2 & 0 & -1 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 1 \\ -1 & 2 & 0 \end{bmatrix}, \\ \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} -1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & -1 & 4 \\ 2 & 1 & 3 \end{bmatrix} \end{matrix}\right\}; \mathfrak{M}_{2\times 3}.$

289 $\left\{\begin{matrix} \begin{bmatrix} 2 & 4 & 3 \\ -1 & 2 & 3 \end{bmatrix}, \begin{bmatrix} -1 & 2 & 1 \\ 2 & 2 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \end{bmatrix}, \\ \begin{bmatrix} 0 & 5 & 3 \\ 2 & 3 & 5 \end{bmatrix}, \begin{bmatrix} -4 & 0 & -1 \\ 5 & 2 & 1 \end{bmatrix}, \begin{bmatrix} -4 & 5 & 2 \\ 7 & 5 & 6 \end{bmatrix} \end{matrix}\right\}; \mathfrak{M}_{2\times 3}.$

290 $\{x, x^2+2, (x-2)^2\}; \mathbf{P}_2.$

291 $\{1, (x+1)^2\}; \mathbf{P}_2.$

292 $\{x, (x+3)^2, (x-3)^2\}; \mathbf{P}_2.$

293 $\{1+x-2x^2, 2x^2-x-2, 4x^2+2x-1\}; \mathbf{P}_2.$

294 $\{1+4x-2x^2, 2-3x+x^2, 3-12x+6x^2\}; \mathbf{P}_2.$

295 $\{2-3x, x^3-4x^2+1, 5x^2-x+3\}; \mathbf{P}_3.$

296 $\{1+x, x-1, x^2, x^3\}; \mathbf{P}_3.$

297 $\{1, x-2, x^2, x^2+x, x^3+1, (x-1)^3, \}; \mathbf{P}_4.$

298 Sea $\mathbf{E}$ un espacio vectorial de dimensión finita n y $S < \mathbf{E}$.

(*a*) Probar que S tiene dimensión finita y que $\dim(S) \leq n$.

(*b*) Demostrar que $S = \mathbf{E}$ si y sólo si $\dim(S) = n$.

En los ejercicios 299 a 306, determinar si el subconjunto S de puntos (x, y, z) en el espacio $\mathbb{R}^3$ que satisface la condición dada es un subespacio de $\mathbb{R}^3$. En caso afirmativo encontrar una base y la dimensión de S.

299 $x = 0.$

300 $x + z = 0.$

301 $x + y + z = 0.$

302 $x+y-z=2$.

303 $x=y$.

304 $x=y=z$.

305 $x^2-z^2=0$.

306 $x+y+z=0$ y $x-y-z=0$.

En los ejercicios 307 a 315, S es el conjunto de polinomios p en el espacio $\mathbf{P}_n$, espacio de polinomios de grado a lo más n, que satisface la condición dada. Determinar si S es subespacio de $\mathbf{P}_n$ y, de ser así, encontrar una base y la dimensión de S.

307 $p'(0)=0$.

308 $p''(0)=0$.

309 $p(0)=p(1)$.

310 $p(0)+p'(0)=0$.

311 $p(0)=p(2)$.

312 p es par ($p(-x)=p(x)\ \forall x$).

313 p es impar ($p(-x)=-p(x)\ \forall x$).

314 p tiene grado a lo más $k<n$ o p es el polinomio constante cero.

315 p tiene grado $k<n$ o p es el polinomio constante cero.

En los ejercicios 316 a 325, encontrar la dimensión del espacio dado en el espacio de funciones $\mathscr{F}(\mathbb{R})$. En los casos donde corresponda, α y β son números reales dados.

316 $\mathrm{gn}(1,e^{\alpha x},e^{\beta x})$, $\alpha\neq\beta$.

317 $\mathrm{gn}(e^{\alpha x},xe^{\alpha x})$.

318 $\mathrm{gn}(1,e^{\alpha x},xe^{\alpha x})$.

319 $\mathrm{gn}(e^{\alpha x},xe^{\alpha x},x^2e^{\alpha x})$.

320 $\mathrm{gn}(\operatorname{senh} x,e^{x},e^{-x})$.

321 $\mathrm{gn}(\operatorname{sen} x,\cos x)$.

322 $\text{gn}(\text{sen}^2 x, \cos^2 x)$.

323 $\text{gn}(\text{senh}^2 x, \cosh^2 x)$.

324 $\text{gn}(1, \cos 2x, \text{sen}^2 2x)$.

325 $\text{gn}(e^{-2x}\cos x, e^{-2x}\,\text{sen}\, x)$.

En los ejercicios 326 a 346, extraer una base del conjunto generador del subespacio dado utilizando el teorema 3.17 (cfr. pág. 164) y determinar la correspondiente dimensión.

326 $\text{gn}((1,-2),(-2,4))$ en $\mathbb{R}^2$.

327 $\text{gn}((3,6),(-1,3))$ en $\mathbb{R}^2$.

328 $\text{gn}((-1,1),(2,1))$ en $\mathbb{R}^2$.

329 $\text{gn}((1,-1,2),(3,-1,1),(1,1,-3),(0,4,-10))$ en $\mathbb{R}^3$.

330 $\text{gn}((2,-1,3),(4,-2,6),(8,-4,12),(-10,5,-15)$ en $\mathbb{R}^3$.

331 $\text{gn}((1,1,1),(-2,1,3),(-1,2,1),(2,2,3))$ en $\mathbb{R}^3$.

332 $\text{gn}((-1,2,3,1),(-2,1,0,2),(-1,1,2,2),(-1,2,0,-2))$ en $\mathbb{R}^4$.

333 $\text{gn}((2,1,3,2),(-1,1,-1,1),(2,7,5,10),(-4,4,-4,4))$ en $\mathbb{R}^4$.

334 $\text{gn}((-1,2,1,3),(-2,3,1,4),(-3,4,2,6),(-1,1,2,3),(1,-6,2,3))$ en $\mathbb{R}^4$.

335 $\text{gn}\left(\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 3 & -1 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} -6 & 6 \\ 2 & -3 \end{bmatrix}\right)$ en $\mathfrak{M}_2$.

336 $\text{gn}\left(\begin{bmatrix} 2 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 4 & 3 \\ 8 & 3 & 5 \end{bmatrix}, \begin{bmatrix} -3 & -6 & -5 \\ -14 & -5 & -9 \end{bmatrix}\right)$ en $\mathfrak{M}_{2\times 3}$.

337 $\text{gn}\left(\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}\right)$ en $\mathfrak{M}_2$.

338 $\text{gn}(x^2-4, x^2+4, 2, 3x-2)$ en $\mathbf{P}$.

339 $\text{gn}(2, 3+2x, 2+3x, x^2+1, x^2-x)$ en $\mathbf{P}$.

340 $\text{gn}(x, x-1, x^2+2, -x-8-3x^2, x^3-x)$ en $\mathbf{P}$.

341 $\text{gn}(2-x, 2+x, x-x^2, -2+2x+3x^2, x^3+1, x^4-x)$ en $\mathbf{P}$.

342 $\text{gn}(1, \cos^2 x, \cos 2x, \text{sen}^2 x)$ en $\mathscr{F}(\mathbb{R})$.

343 $\text{gn}(\text{sen}\,4x, \text{sen}\,2x, \text{sen}\,3x, \cos x)$ en $\mathscr{F}(\mathbb{R})$.

344 $\text{gn}(\cos 8x, \cos(2x), \cos 5x, \cos 3x)$ en $\mathscr{F}(\mathbb{R})$.

345 $\text{gn}(1, e^x, e^{-x}, \cosh x, \text{senh}\,x)$ en $\mathscr{F}(\mathbb{R})$.

346 $\text{gn}\left(1, \cos x, \cos 2x, \dfrac{\text{sen}(\frac{5}{2}x)}{2\,\text{sen}(\frac{1}{2}x)}\right)$ en $\mathscr{F}((0, 2\pi))$.

347 Sea $\mathbf{P}_n$ el espacio de polinomios. A cada polinomio $p(x) = \sum_{k=0}^{n} a_0 x^k$ se le asocia el vector $[p] = (a_0, a_1, \ldots, a_n)$ de $\mathbb{R}^{n+1}$. Probar:

(a) $[p] = 0 \Leftrightarrow p$ es el polinomio constante cero $(p = \theta)$.

(b) $[p+q] = [p] + [q]$ para todo par $p, q \in \mathbf{P}_n$.

(c) $[\alpha p] = \alpha[p]$ para todo $\alpha \in \mathbb{R}$ y para todo $p \in \mathbf{P}_n$.

(d) Los polinomios p_i, $i = 1, \ldots, m$, son L.I. en $\mathbf{P}_n$ si y sólo si los vectores asociados correspondientes $[p_i]$ son L.I. en $\mathbb{R}^{n+1}$.

En los ejercicios 348 a 351, utilice el último inciso del ejercicio precedente para encontrar una base y la dimensión del subespacio dado en el espacio de polinomios.

348 $\text{gn}(1 + 4x - 2x^2 + x^3, 1 - 9x + 3x^2 - 2x^3, 5 + 7x - 5x^2 + 3x^3, 5 - 6x - x^3)$.

349 $\text{gn}(1 - x + 2x^2 + x^3, -2 + 3x + 2x^2 + 2x^3, 6 - 8x - 2x^3, -1 + 2x + x^2 + 3x^3)$.

350 $\text{gn}(2 + x + x^2 - 3x^3, 5 - 2x + 7x^2 - 9x^3, -1 - 2x + x^2 + x^3, -1 + 2x^2)$.

351 $\text{gn}(3 - x + 2x^2 + x^3, 1 + 4x + 5x^2 + 7x^3, -1 + 2x + x^2 + 3x^3, 2 - x + x^2 - x^3)$.

352 El conjunto de matrices triangulares superiormente de orden n, S, es un subespacio vectorial del espacio $\mathfrak{M}_n$ (cfr. ejercicio propuesto 110 de este capítulo).

(a) Encontrar una base y la dimensión de este subespacio para los casos $n = 2$ y $n = 3$.

(b) Hallar una base y $\dim(S)$ en $\mathfrak{M}_n$ para cualquier n.

353 El conjunto de matrices simétricas de orden n, S, es un subespacio de $\mathfrak{M}_n$ (cfr. ejemplo 3.25).

(a) Hallar una base y $\dim(S)$ en el caso $n = 3$.

(b) Encontrar una base y $\dim(S)$ en $\mathfrak{M}_n$ para todo n.

354 Una matriz $A \in \mathfrak{M}_n$ es antisimétrica si $A^t = -A$. Probar que el conjunto S de todas estas matrices, las antisimétricas, es un subespacio de $\mathfrak{M}_n$, encontrar una base para S y calcular $\dim(S)$.

355 Sea S el conjunto de matrices cuadradas de orden n cuya traza es cero. Probar que S es un subespacio de $\mathfrak{M}_n$, encontrar una base para S y $\dim(S)$. (Si $A = [a_{ij}]$, $\text{tra}(A) = \sum_{i=1}^{n} a_{ii}$.)

356 Sea S un subespacio de un espacio vectorial $\mathbf{E}$ con $\dim(\mathbf{E}) = n$. Encontrar una base y la dimensión del espacio cociente $\mathbf{E}/S$ (cfr. ejercicio resuelto 29).

357 Sea S el subespacio de $\mathbb{R}^\infty$ que consiste en todas las sucesiones (a_n) tales que $a_n = 0$ salvo un número finito de índices n. Probar que $S < \mathbb{R}^\infty$, encontrar una base y la dimensión de S.

358 Sean $n > 1$ un número entero y $c_k \in \mathbb{R}$, $k = 0, 1, \ldots, n$, escalares distintos entre sí. Para cada $k = 0, 1, \ldots, n$ se define

$$L_k(x) = \prod_{\substack{i=1 \\ i \neq k}}^{n} \frac{x - c_i}{c_k - c_i}.$$

(a) Probar que

$$L_k(c_i) = \begin{cases} 0 \text{ si } i \neq k \\ 1 \text{ si } i = k \end{cases}$$

para todo k y para todo i.

(b) Mostrar que $L_k(x)$ es un polinomio de grado a lo más n. Los polinomios $L_k(x)$ se llaman polinomios de Lagrange (asociados a los escalares c_k).

(c) Demostrar que los polinomios de Lagrange $L_0, L_1, \ldots, L_n$ son L.I. en $\mathbf{P}_n$; por ende, ya que $\dim(\mathbf{P}_n) = n + 1$, forman una base de $\mathbf{P}_n$.

En los ejercicios 359 a 370, completar el conjunto L.I., $\mathscr{B}$, a una base del espacio indicado.

359 $\mathscr{B} = \{(-1, 2, 3), (2, -1, 2)\}$; $\mathbb{R}^3$.

360 $\mathscr{B} = \{(1, -1, 2), (-1, 2, 2)\}$; $\mathbb{R}^3$.

361 $\mathscr{B} = \{(2, -1, 2, 2), (-1, 2, 3, -2)\}$; $\mathbb{R}^4$.

362 $\mathscr{B} = \{(2, -2, 3, 1), (-2, 0, 3, 1)\}$; $\mathbb{R}^4$.

363 $\mathscr{B} = \{(-1, 0, 2, 3, 2), (-1, 1, 2, 3, -2), (-1, 1, 0, 1, 1)\}$; $\mathbb{R}^5$.

364 $\mathscr{B} = \{(-1, 2, 1, 4, 3), (-2, 1, 4, 2, 3)\}$; $\mathbb{R}^5$.

365 $\mathscr{B} = \left\{ \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix} \right\}$; $\mathfrak{M}_2$.

366 $\mathscr{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \right\}$; $\mathfrak{M}_2$.

367 $\mathscr{B} = \left\{ \begin{bmatrix} -1 & 2 & 1 \\ 2 & -1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} -2 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \right\}$; $\mathfrak{M}_{2\times 3}$.

368 $\mathscr{B} = \{x - 1, x^2 + 1\}$; $\mathbf{P}_3$.

369 $\mathscr{B} = \{2, x-2, (x-2)^2\}$; $\mathbf{P}_4$.

370 $\mathscr{B} = \{x-1, (x-1)^2, (x-1)^3, (x-1)^4\}$; $\mathbf{P}_5$.

371 Sean $\mathbf{E}$ un espacio vectorial de dimensión finita y S_1, S_2 subespacios de $\mathbf{E}$. Si $\mathbf{E} = S_1 \oplus S_2$, $\mathscr{B}_1 = \{\vec{u}_1, \ldots, \vec{u}_k\}$ es una base de S_1 y $\mathscr{B} = \{\vec{v}_1, \ldots, \vec{v}_l\}$ es una base de S_2, mostrar que $\mathscr{B}_1 \cap \mathscr{B}_2 = \emptyset$ y que $\mathscr{B}_1 \cup \mathscr{B}_2$ es una base de $\mathbf{E}$ (cfr. ejercicio resuelto 27).

372 Sean $\mathbf{E}$ un espacio vectorial de dimensión finita y S_1, S_2 subespacios de $\mathbf{E}$. Si $\mathscr{B}_1 = \{\vec{u}_1, \ldots, \vec{u}_k\}$ es una base de S_1, $\mathscr{B}_2 = \{\vec{v}_1, \ldots, \vec{v}_l\}$ es una base de S_2, $\mathscr{B}_1 \cap \mathscr{B}_2 = \emptyset$ y $\mathscr{B}_1 \cup \mathscr{B}_2$ es una base de $\mathbf{E}$, probar que $\mathbf{E} = S_1 \oplus S_2$.

373 Sean $\mathbf{E}$ un espacio vectorial de dimensión finita y $S_1 < \mathbf{E}$. Probar que existe $S_2 < \mathbf{E}$ tal que $\mathbf{E} = S_1 \oplus S_2$.

374 Probar que un espacio vectorial tiene dimensión infinita si y sólo si contiene un conjunto infinito linealmente independiente.

375 Sean $\alpha, \beta \in \mathbb{R}$ un par de números reales con $\beta \neq 0$. Sea S el conjunto de sucesiones de números reales que satisfacen la relación

$$x_{n+2} + \alpha x_{n+1} + \beta x_n = 0 \; \forall n.$$

En el ejercicio resuelto 44 se probó que $S < \mathbb{R}^\infty$.

(a) Mostrar que si $\lambda \in \mathbb{R}$ es raíz doble del polinomio $X^2 + \alpha X + \beta = 0$, entonces las sucesiones $u = (u_n)$ y $v = (v_n)$, con $u_n = \lambda^n$ y $v_n = n\lambda^n$, pertenecen a S.

(b) Demostrar que $\{u, v\}$ es una base de S.

376 Sean S_1 el conjunto de matrices de la forma

$$\begin{bmatrix} a & b \\ c & a \end{bmatrix}, \; a, b, c \in \mathbb{R}$$

y S_2 el conjunto de matrices de la forma

$$\begin{bmatrix} 0 & a \\ -a & b \end{bmatrix}, \; a, b \in \mathbb{R}.$$

Probar que $S_1, S_2 < \mathfrak{M}_2$; encontrar bases y dimensiones de S_1, S_2, $S_1 + S_2$ y $S_1 \cap S_2$.

En los problemas 377 a 382, hallar una base y la dimensión del espacio solución del sistema homogéneo indicado.

377

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= 0 \\ x_1 - 5x_2 + 3x_3 &= 0 \\ 2x_1 - x_2 + 3x_3 &= 0. \end{aligned}$$

378

$$\begin{aligned} x_1 - x_2 + x_3 + 2x_4 - x_5 &= 0 \\ 2x_1 - x_2 + 3x_3 - x_4 + 2x_5 &= 0 \\ -3x_1 + 2x_2 - 3x_3 + 2x_4 - x_5 &= 0. \end{aligned}$$

379
$$\begin{aligned} -x_1+3x_2-x_3+x_4 &= 0\\ 2x_1-3x_2+x_3-5x_4 &= 0. \end{aligned}$$

380
$$\begin{aligned} x_1+3x_2-3x_3+x_4 &= 0\\ 2x_1-x_2+3x_3-x_4 &= 0\\ -x_1+x_2-3x_3+x_4 &= 0\\ 3x_1-x_2+x_3-x_4 &= 0. \end{aligned}$$

381
$$\begin{aligned} -2x_1+3x_2-x_3+7x_4-2x_5+3x_6 &= 0\\ -x_1+x_2-x_3+4x_4+3x_5-2x_6 &= 0\\ 3x_1-x_2+4x_3-5x_4+3x_5-6x_6 &= 0. \end{aligned}$$

382
$$\begin{aligned} x_1+2x_2-3x_3+x_4-3x_5+2x_6 &= 0\\ -2x_1-x_2-2x_3+3x_4-3x_5-x_6 &= 0\\ 3x_1-2x_2+x_3+5x_4-3x_5+x_6 &= 0\\ 4x_1-x_2+x_3-4x_4+2x_5+x_6 &= 0. \end{aligned}$$

En los ejercicios 383 a 388, encontrar bases para el espacio fila, el espacio columna, el espacio nulo, las respectivas dimensiones y el rango de la matriz indicada.

383
$$\begin{bmatrix} 1 & -1 & 1\\ 2 & -1 & 3\\ 3 & -2 & 4 \end{bmatrix}.$$

384
$$\begin{bmatrix} 1 & -2 & 3\\ -1 & 3 & 1\\ 4 & -1 & 2\\ 0 & -1 & 1 \end{bmatrix}.$$

385
$$\begin{bmatrix} -2 & -1 & 3 & 1\\ 3 & 4 & -1 & 1\\ 2 & 1 & 1 & 2 \end{bmatrix}.$$

386
$$\begin{bmatrix} 1 & -1 & 2 & 3\\ -1 & 0 & 1 & 2\\ -1 & 1 & 2 & 1\\ 1 & -1 & 6 & 7 \end{bmatrix}.$$

387
$$\begin{bmatrix} 4 & -2 & 1 & 0 & 3 & 1\\ -2 & 3 & -1 & 1 & 1 & 0\\ 0 & 2 & -1 & 2 & 1 & 1\\ 1 & -1 & 1 & 1 & -1 & 2 \end{bmatrix}.$$

388
$$\begin{bmatrix} 1 & -1 & 2 & 1 & 3 & 5\\ 2 & -1 & 0 & 4 & 6 & 2\\ -2 & 3 & 1 & 0 & 2 & 0\\ 2 & 2 & 1 & 3 & -1 & 2\\ 1 & -1 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

389 Encontrar una ecuación para el subespacio S generado por los vectores $(-1,2,-1,1,0)$, $(2,-1,1,2,-3)$, $(2,-2,4,1,1)$ de $\mathbb{R}^5$; es decir, hallar un sistema homogéneo $A\vec{x}=\vec{0}$ tal que el espacio solución sea S (cfr. ejercicio resuelto 46).

390 Encontrar una ecuación para el subespacio S generado por los vectores $(1,-3,2,1,1)$, $(1,-1,3,2,1)$, $(-2,1,-2,-1,1)$ de $\mathbb{R}^5$; es decir, hallar un sistema homogéneo $A\vec{x}=\vec{0}$ tal que el espacio solución sea S.

391 Sean $S_1=\{(x_1,x_2,x_3,x_4)\,|\,x_2-x_3+2x_4=0\}$, $S_2=\{(x_1,x_2,x_3,x_4)\,|\,x_1=x_4,\,x_3=-2x_4\}$. Encontrar bases de S_1+S_2, $S_1\cap S_2$ y las respectivas dimensiones.

392 Pruebe que si S_1 y S_2 son subespacios de $\mathbb{R}^3$ con $\dim(S_1)=\dim(S_2)=2$, entonces $S_1\cap S_2\neq\{\vec{0}_{\mathbb{R}^3}\}$.

393 Sean S_1 y S_2 los subespacios de $\mathbb{R}^5$ generados por

$$\{(-1,2,-1,2,2),(2,-1,2,3,1),(3,-1,2,1,0)\}$$

y

$$\{(-2,3,1,2,1),(-1,1,1,2,1),(2,-4,2,1,3)\},$$

respectivamente. Hallar las bases y las dimensiones de los espacios S_1+S_2 y $S_1\cap S_2$.

394 Sean S_1 y S_2 los subespacios de $\mathbb{R}^5$ generados por

$$\{(1,-2,3,-2,1),(-2,1,2,-3,4),(2,1,-2,2,1)\}$$

y

$$\{(1,-3,-1,1,1),(1,-2,-1,1,2),(3,4,-1,2,1)\},$$

respectivamente. Hallar las bases y las dimensiones de los espacios S_1+S_2 y $S_1\cap S_2$.

395 Sean S_1 y S_2 los subespacios de $\mathbb{R}^5$ generados por

$$\{(1,-2,1,3),(-1,-2,3,2)\}\text{ y }\{(1,2,1,-2),(1,1,-2,1)\},$$

respectivamente. Hallar las bases y las dimensiones de los espacios S_1+S_2 y $S_1\cap S_2$.

396 Sean S_1 y S_2 los subespacios de $\mathbf{P}$ generados por $\{1-x+x^2+2x^3,2-3x+2x^2-x^3,-1+2x+3x^2-4x^3\}$ y $\{(-2+3x-4x^2+x^3,5-x+3x^2+2x^3,-4+x+2x^2-5x^3\}$, respectivamente. Hallar las bases y las dimensiones de los espacios S_1+S_2 y $S_1\cap S_2$.

397 Sean S_1 y S_2 los subespacios de $\mathbf{P}$ generados por $\{x^2+2x^3,2-5x-x^2-x^3,-1+3x^2-4x^3\}$ y $\{(3x-5x^2-3x^3,2-3x-x^2-x^3,1+x+x^2-7x^3\}$, respectivamente. Hallar las bases y las dimensiones de los espacios S_1+S_2 y $S_1\cap S_2$.

Espacios vectoriales complejos (respuestas en página 1080)

398 Sea n un entero mayor o igual a 1. Probar que el conjunto $\mathbb{C}^n$ de vectores $(z_1,\ldots,z_n)$, $z_i \in \mathbb{C}$ para cada $i = 1,\ldots,n$, con las operaciones:

$$\begin{aligned}(z_1,\ldots,z_n)+(w_1,\ldots,w_n) &= (z_1+w_1,\ldots,z_n+w_n)\\ \alpha(z_1,\ldots,z_n) &= (\alpha z_1,\ldots,\alpha z_n),\end{aligned}$$

donde $z_i + w_i$ y αz_i son las operaciones suma y multiplicación de números complejos, es un espacio vectorial sobre $\mathbb{C}$.

399 Sean n y m un par de números enteros positivos. Probar que $\mathfrak{M}_{m\times n}(\mathbb{C})$, el conjunto de matrices $[a_{ij}]$ con componentes complejas, es un espacio vectorial con las operaciones:

$$\begin{aligned}[a_{ij}]+[b_{ij}] &= [a_{ij}+b_{ij}]\\ \alpha[a_{ij}] &= [\alpha a_{ij}],\end{aligned}$$

donde $a_{ij}+b_{ij}$ y αa_{ij} son la suma y el producto de números complejos, es un espacio vectorial sobre $\mathbb{C}$.

400 Sea $\mathbf{P}(\mathbb{C})$ el conjunto de polinomios con coeficientes en $\mathbb{C}$. Mostrar que $\mathbf{P}(\mathbb{C})$ con la suma de polinomios y la multiplicación de un escalar por un polinomio usuales (pero sobre el campo $\mathbb{C}$) es un espacio vectorial complejo.

401 Probar que $\mathscr{F}(X,\mathbb{C}^n)$, el conjunto de funciones $f : X \to \mathbb{C}^n$, con las operaciones usuales de suma de funciones y multiplicación de un escalar por una función, es un espacio vectorial sobre $\mathbb{C}$.

402 Sea $\mathbb{C}^\infty$ el conjunto de sucesiones (a_n), con $a_n \in \mathbb{C}$ para todo $n \in \mathbb{N}$. Se definen, para cada $(a_n),(b_n) \in \mathbb{C}^\infty$ y para cada $\alpha \in \mathbb{C}$:

$$\begin{aligned}(a_n)+(b_n) &= (a_n+b_n)\\ \alpha(a_n) &= (\alpha a_n).\end{aligned}$$

Mostrar que $\mathbb{C}^\infty$, junto con estas operaciones, es un espacio vectorial complejo.

403 Sea $\mathbf{E}$ el conjunto de n-adas ordenadas $(z_1,\ldots,z_n)$, con los $z_i \in \mathbb{C}$. Se definen, para cada par $(z_1,\ldots,z_n)$, $(w_1,\ldots,w_n) \in \mathbf{E}$ y para cada $\alpha \in \mathbb{R}$, las operaciones

$$\begin{aligned}(z_1,\ldots,z_n)+(w_1,\ldots,w_n) &= (z_1+w_1,\ldots,z_n+w_n)\\ \alpha(z_1,\ldots,z_n) &= (\alpha z_1,\ldots,\alpha z_n),\end{aligned}$$

donde $z_i + w_i$ es la suma de números complejos y αz_i es la multiplicación de un número real por un complejo $(\alpha(a+bi) = (\alpha a)+(\alpha b)i)$. Probar que $\mathbf{E}$ junto con estas operaciones es un espacio vectorial real; es decir, sobre $\mathbb{R}$.

404 Muestre que si en $\mathfrak{M}_{m\times n}(\mathbb{C})$ se restringe la operación multiplicación por un escalar a números reales el espacio resultante $\mathbf{E}$ es un espacio vectorial real. Determinar una base y la dimensión de este espacio.

405 Encontrar una base y la dimensión de

$$\text{gn}((1,-1+i,2i,3+i),(i,-i,2i,1+4i),(-i,-2+i,-4-6i,-5-6i)) \text{ en } \mathbb{C}^4.$$

406 Completar el conjunto L.I. $\{(1,-2i,2i),(1,-1,2)\}$ a una base de $\mathbb{C}^3$.

407 Completar el conjunto L.I. $\{(1,i,-i,2,0),(1,-1,1,1,i),(2,-1,i,-i,1)\}$ a una base de $\mathbb{C}^5$.

408 Encontrar una base y la dimensión de

$$\operatorname{gn}\left(\begin{bmatrix} 1 & -i & 1 \\ 2 & -1 & 1 \end{bmatrix},\begin{bmatrix} -2 & 2 & -i \\ 1 & 0 & i \end{bmatrix},\begin{bmatrix} -2+i & 3 & 0 \\ 1+2i & -i & 2i \end{bmatrix}\right)$$

en $\mathfrak{M}_{2\times 3}(\mathbb{C})$.

409 Completar el conjunto L.I.

$$\left\{\begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix},\begin{bmatrix} 1 & -1 \\ 2 & -i \end{bmatrix}\right\}$$

a una base de $\mathfrak{M}_2(\mathbb{C})$.

410 Sean $\alpha,\beta\in\mathbb{R}$ un par de números reales con $\beta\neq 0$. Sea S el conjunto de sucesiones en $\mathbb{C}^\infty$ (cfr. ejercicio 402) que satisfacen la relación

$$x_{n+2}+\alpha x_{n+1}+\beta x_n = 0\ \forall n.$$

(cfr. ejercicio resuelto 44).

(a) Probar que S es un subespacio de $\mathbb{C}^\infty$.

(b) Mostrar que si $\lambda,\mu\in\mathbb{C}$ son raíces complejas (conjugadas) del polinomio $X^2+\alpha X+\beta$, entonces las sucesiones $u=(u_n)$ y $v=(v_n)$, con $u_n=\lambda^n$ y $v_n=\mu^n$, pertenecen a S.

(c) Demostrar que $\{u,v\}$ es una base de S.

4 Espacios con producto interior y espacios normados

El propósito de este capítulo es dotar de una "geometría" análoga a la de $\mathbb{R}^2$ y $\mathbb{R}^3$ a otro tipo de espacios vectoriales, de manera que, conceptos tales como producto punto, ángulo entre vectores, perpendicularidad, norma de un vector y distancia entre vectores adquieran significado en otros espacios; por ejemplo, en el de polinomios, en el espacio de funciones, en el espacio de matrices o en el de sucesiones; aun si en éstos no existe una forma física de "observar" dicha geometría. Lo anterior, como veremos después, no es un simple capricho con fines de generalización, sino que es de gran importancia teórica y aplicada. Para poder hacerlo, es necesario generalizar el producto punto de $\mathbb{R}^n$ a estos espacios, ya que a partir de él, como vimos antes, fue posible precisar dichos conceptos geométricos en el espacio $\mathbb{R}^n$. A esto nos abocaremos en la primera sección de este capítulo; para que en las siguientes secciones y subsecciones podamos introducir los conceptos geométricos de los que hemos hablado. Como siempre, los últimos dos segmentos están dedicados a la resolución de ejercicios y problemas, y a proponer ejercicios para que el lector practique.

4.1 Espacios con producto interior

Para lograr extender el concepto de producto punto de los espacios $\mathbb{R}^n$ a otros espacios vectoriales necesitamos abstraer las propiedades esenciales del producto punto que son, precisamente, las que enunciamos en la página 123. Recordemos estas características esenciales.[1]

En $\mathbb{R}^n$ el *producto interior* (producto punto) se define como:

$$(x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = \sum_{i=1}^{n} x_i y_i$$

y tiene las propiedades:

1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$,
2. $(\lambda \vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v})$,
3. $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$,
4. $\vec{u} \cdot \vec{u} \geq 0$,
5. $\vec{u} \cdot \vec{u} = 0 \Leftrightarrow \vec{u} = \vec{0}_{\mathbb{R}^n}$;

$\forall \vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$; $\forall \lambda \in \mathbb{R}$.

[1] Les llamamos propiedades esenciales porque cualquier otra propiedad del producto punto se deduce de éstas.

4.1.1 Definiciones, ejemplos y propiedades

Definición 4.1 ***(Definición de producto interior)*** *Sea* **E** *un espacio vectorial. Un* ***producto interior*** *(****producto escalar****) en* **E** *es una función que a cada par de vectores* $\vec{u}, \vec{v} \in \mathbf{E}$*, le asigna un número real denotado como* $\langle \vec{u}, \vec{v} \rangle$*, llamado* el producto interior de $\vec{u}$ *con* $\vec{v}$ *(también se acostumbra denotarlo por* $(\vec{u} \,|\, \vec{v})$*), que satisface las siguientes condiciones:*

1. $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$ (Simetría)
2. $\langle \lambda\vec{u}, \vec{v} \rangle = \lambda\langle \vec{u}, \vec{v} \rangle$ (Homogeneidad)
3. $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$ (Aditividad)
4. $\langle \vec{u}, \vec{u} \rangle \geq 0$ (Positividad)[2]
5. $\langle \vec{u}, \vec{u} \rangle = 0 \Leftrightarrow \vec{u} = \vec{0}_E$

$\forall \vec{v}, \vec{u}, \vec{w} \in \mathbf{E}$ *y* $\forall \lambda \in \mathbb{R}$.

► **Ejemplo 4.1** En $\mathbb{R}^2$, si $\vec{x} = (x_1, x_2), \vec{y} = (y_1, y_2)$ se define

$$\langle \vec{x}, \vec{y} \rangle = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2.$$

Demostrar que $\langle \vec{x}, \vec{y} \rangle$ es un producto interior en $\mathbb{R}^2$. ◄

DEMOSTRACIÓN ■ Si $\vec{u} = (a,b), \vec{v} = (c,d), \vec{w} = (e,f) \in \mathbb{R}^2$ y $\lambda \in \mathbb{R}$:

1.
$$\langle \vec{u}, \vec{v} \rangle = ac - 2ad - 2bc + 5bd$$

y

$$\begin{aligned} \langle \vec{v}, \vec{u} \rangle &= \langle (c,d), (a,b) \rangle \\ &= ca - 2cb - 2da + 5db \\ &= ac - 2ad - 2bc + 5bd. \end{aligned}$$

Es decir, $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$.

2.
$$\begin{aligned} \langle \lambda\vec{u}, \vec{v} \rangle &= \langle (\lambda a, \lambda b), (c,d) \rangle \\ &= \lambda ac - 2\lambda ad - 2\lambda bc + 5\lambda bd \\ &= \lambda[ac - 2ad - 2bc + 5bd] \\ &= \lambda\langle \vec{u}, \vec{v} \rangle \end{aligned}$$

3.
$$\begin{aligned} \langle \vec{u}, \vec{v} + \vec{w} \rangle &= \langle (a,b), (c+e, d+f) \rangle \\ &= a(c+e) - 2a(d+f) - 2b(c+e) + 5b(d+f). \\ \langle \vec{u}, \vec{v} \rangle &= ac - 2ad - 2bc + 5bd, \\ \langle \vec{u}, \vec{w} \rangle &= ae - 2af - 2be + 5bf; \quad \therefore \\ \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle &= a(c+e) - 2a(d+f) - 2b(c+e) + 5b(d+f); \end{aligned}$$

[2] Con frecuencia, a las propiedades 2 y 3, se les cita diciendo que el producto interior es bilineal; y a las propiedades 4 y 5, mencionando que el producto interior es definido positivo.

esto es,

$$\langle\vec{u},\vec{v}+\vec{w}\rangle=\langle\vec{u},\vec{v}\rangle+\langle\vec{u},\vec{w}\rangle.$$

4.
$$\begin{aligned}\langle\vec{u},\vec{u}\rangle &= a^2-2ab-2ba+5b^2\\ &= a^2-4ab+5b^2\\ &= a^2-4ab+4b^2+b^2\\ &= (a-2b)^2+b^2\geq 0;\end{aligned}$$

de donde, $\langle\vec{u},\vec{u}\rangle\geq 0$.

5. Si $\vec{u}=(0,0)$, claramente $\langle\vec{u},\vec{u}\rangle=0$. Si $\vec{u}=(a,b)$ y $\langle\vec{u},\vec{u}\rangle=0$, entonces $(a-2b)^2+b^2=0$, de ahí que $a=b=0$ y que $\vec{u}=(a,b)=(0,0)$. ■

► **Ejemplo 4.2** En $\mathbb{R}^n$, el producto punto de vectores es un producto interior.◄

► **Ejemplo 4.3** *(Producto interior en el espacio de funciones continuas).* En el espacio de las funciones continuas, $C[a,b]$ se define, para $f,g\in C[a,b]$,

$$\langle f,g\rangle=\int_a^b f(t)g(t)dt.$$

Comprobemos que $\langle f,g\rangle$ es un producto interior:

1. Si $f,g\in C[a,b]$

$$\langle f,g\rangle=\int_a^b f(t)g(t)dt=\int_a^b g(t)f(t)dt=\langle g,f\rangle.$$

2. Si $\lambda\in\mathbb{R}$ y $f,g\in C[a,b]$,

$$\begin{aligned}\langle\lambda f,g\rangle &= \int_a^b(\lambda f)(t)g(t)dt\\ &= \int_a^b\lambda f(t)g(t)dt\\ &= \lambda\int_a^b f(t)g(t)dt\\ &= \lambda\langle f,g\rangle.\end{aligned}$$

3. Si $f,g,h\in C[a,b]$

$$\begin{aligned}\langle f,g+h\rangle &= \int_a^b f(t)(g+h)(t)dt\\ &= \int_a^b f(t)[g(t)+h(t)]dt\\ &= \int_a^b [f(t)g(t)+f(t)h(t)]dt\end{aligned}$$

$$
\begin{aligned}
&= \int_a^b f(t)g(t)dt + \int_a^b f(t)h(t)dt \\
&= \langle f,g\rangle + \langle f,h\rangle.
\end{aligned}
$$

4. Si $f \in C[a,b]$

$$
\begin{aligned}
\langle f,f\rangle &= \int_a^b f(t)f(t)dt \\
&= \int_a^b [f(t)]^2 dt \geq 0.
\end{aligned}
$$

5. Si $f = \theta, \theta(x) = 0 \,\forall x \in [a,b]$, entonces

$$
\begin{aligned}
\langle f,f\rangle &= \langle \theta,\theta\rangle \\
&= \int_a^b \theta^2(t)dt \\
&= \int_a^b 0\cdot dt \\
&= 0.
\end{aligned}
$$

Si $f \in C[a,b]$ y $\langle f,f\rangle = 0$, es decir $\int_a^b f^2(t)dt = 0$, sea

$$F(x) = \int_a^x f^2(t)dt.$$

Entonces,

$$
\begin{aligned}
F(a) &= \int_a^a f^2(t)dt = 0, \\
F(b) &= \int_a^b f^2(t)dt = 0.
\end{aligned}
$$

Ya que f es continua, f^2 también lo es. Así, por el teorema fundamental del cálculo,[3]

$$F'(x) = f^2(x) \geq 0$$

$\forall x \in [a,b]$. Luego, F es creciente en $[a,b]$ (pues su derivada es positiva ahí). Entonces,

$$F(a) \leq F(x) \leq F(b),$$

$\forall x \in [a,b]$; y puesto que $F(a) = 0 = F(b)$, se tiene que $F(x) = 0 \;\forall x \in [a,b]$. Por tanto, $f^2(x) = F'(x) = 0 \,\forall x$ y, por ende,

$$f(x) = 0 \,\forall x \in [a,b].$$

Es decir, f es la función θ, el neutro aditivo de $C[a,b]$.◄

[3]Recuerde que el teorema fundamental del cálculo establece que si $g \in C[a,b]$ y se define $G(x) = \int_a^x g(t)\,dt$, $x \in [a,b]$, entonces $G'(x) = g(x) \;\forall x \in [a,b]$.

▶ **Ejemplo 4.4** Calcular $\langle f,g\rangle$ en $C[0,1]$, si $f(x)=x$ y $g(x)=e^x$.◀

Solución Haciendo $u=x$, $dv=e^x dx$ e integrando por partes, se tiene

$$\begin{aligned}\langle f,g\rangle &= \int_0^1 xe^x dx\\ &= xe^x|_0^1 - \int_0^1 e^x dx\\ &= e - e^x|_0^1\\ &= e-(e-1)\\ &= 1. \quad ✔\end{aligned}$$

El producto interior definido en $C[a,b]$ se puede motivar intuitivamente de la manera siguiente: sean N un entero no negativo, $h=\dfrac{b-a}{N}$ y $x_j=a+jh$, $j=0,1,\ldots,N$, la partición correspondiente al dividir el intervalo $[a,b]$ en N subintervalos cada uno de longitud h. Sea $x_k^*\in[x_{k-1},x_k]$ un punto cualquiera de este k-ésimo subintervalo, para cada $k=1,2,\ldots,N$. Entonces, si definimos

$$\begin{aligned}a_k &= f(x_k^*)\sqrt{h} \quad \text{y}\\ b_k &= g(x_k^*)\sqrt{h}\end{aligned}$$

para cada $k=1,2,\ldots,N$, tenemos, por definición de la integral definida, que para N grande

$$\begin{aligned}\int_a^b f(x)g(x)dx &\approx \sum_{k=1}^N f(x_k^*)g(x_k^*)h\\ &= (a_1,a_2,\ldots,a_N)\cdot(b_1,b_2,\ldots,b_N).\end{aligned}$$

De hecho, la aproximación es exacta en el límite; es decir,

$$\begin{aligned}\int_a^b f(x)g(x)dx &= \lim_{N\to\infty}\sum_{k=1}^N f(x_k^*)g(x_k^*)h\\ &= \lim_{N\to\infty}(a_1,a_2,\ldots,a_N)\cdot(b_1,b_2,\ldots,b_N),\end{aligned}$$

lo cual significa que el producto interior en $C[a,b]$ es, en este sentido, un caso límite del producto punto de estos vectores en $\mathbb{R}^N$ cuando N tiende a infinito.

▶ **Ejemplo 4.5** Sea $\mathfrak{M}_{2\times 2}$ el espacio vectorial de las matrices cuadradas de orden 2. Se define para $A=[a_{ij}]$, $B=[b_{ij}]\in\mathfrak{M}_{2\times 2}$,

$$\langle A,B\rangle = a_{11}b_{11}+a_{12}b_{12}+a_{21}b_{21}+a_{22}b_{22}.$$

Por ejemplo,

$$\left\langle \begin{bmatrix}-1 & 1\\ 2 & 1\end{bmatrix}, \begin{bmatrix}3 & 0\\ -2 & 2\end{bmatrix}\right\rangle = (-1)(3)+(1)(0)+(2)(-2)+(1)(2) = -5.$$

Demostrar que $\langle A,B\rangle$ es un producto interior en $\mathfrak{M}_{2\times 2}$.◀

DEMOSTRACIÓN 1.

$$\begin{aligned}\langle A,B\rangle &= a_{11}b_{11}+a_{12}b_{12}+a_{21}b_{21}+a_{22}b_{22}\\ &= b_{11}a_{11}+b_{12}a_{12}+b_{21}a_{21}+b_{22}a_{22}\\ &= \langle B,A\rangle .\end{aligned}$$

2. Si $\lambda\in\mathbb{R}$ y $A=[a_{ij}]\in\mathfrak{M}_{2\times 2}$, entonces

$$\begin{aligned}\langle \lambda A,B\rangle &= (\lambda a_{11})b_{11}+(\lambda a_{12})b_{12}+(\lambda a_{21})b_{21}+(\lambda a_{22})b_{22}\\ &= \lambda(a_{11}b_{11}+a_{12}b_{12}+a_{21}b_{21}+a_{22}b_{22})\\ &= \lambda\langle A,B\rangle .\end{aligned}$$

3. Si $A=[a_{ij}]$,$B=[b_{ij}]$,$C=[c_{ij}]\in\mathfrak{M}_{2\times 2}$, entonces,

$$\begin{aligned}\langle A,B+C\rangle &= a_{11}(b_{11}+c_{11})+a_{12}(b_{12}+c_{12})\\ &\quad +a_{21}(b_{21}+c_{21})+a_{22}(b_{22}+c_{22})\\ &= (a_{11}b_{11}+a_{12}b_{12}+a_{21}b_{21}+a_{22}b_{22})\\ &\quad +(a_{11}c_{11}+a_{12}c_{12}+a_{21}c_{21}+a_{22}c_{22})\\ &= \langle A,B\rangle+\langle A,C\rangle .\end{aligned}$$

4. Si $A=[a_{ij}]\in\mathfrak{M}_{2\times 2}$, entonces,

$$\langle A,A\rangle = a_{11}^2+a_{12}^2+a_{21}^2+a_{22}^2\geq 0.$$

5. Si $A=\begin{bmatrix}0&0\\0&0\end{bmatrix}$, entonces claramente $\langle A,A\rangle=0$. Inversamente supongamos que

$$A=[a_{ij}]\in\mathfrak{M}_{2\times 2}$$

y que $\langle A,A\rangle=0$, entonces,

$$a_{11}^2+a_{12}^2+a_{21}^2+a_{22}^2=0;$$

lo cual implica $a_{11}=a_{12}=a_{21}=a_{22}=0$ y, por tanto, $A=\begin{bmatrix}0&0\\0&0\end{bmatrix}$.

De 1, 2, 3, 4 y 5 $\langle A,B\rangle$ es un producto interior en $\mathfrak{M}_{2\times 2}$. ■

► **Ejemplo 4.6** *(Producto interior en matrices).* En general, en el espacio vectorial $\mathfrak{M}_{m\times n}$, si $A=[a_{ij}],B=[b_{ij}]$ y se define

$$\langle A,B\rangle=\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_{ij}b_{ij}\right),$$

entonces, $\langle A,B\rangle$ es un producto interior. En realidad, si se reflexiona un poco, es inmediato ver que si $\vec{F}_i$ y $\vec{G}_i$ son, respectivamente, las filas i de A y B, consideradas como vectores de $\mathbb{R}^n$, entonces,

$$\langle A,B\rangle=\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_{ij}b_{ij}\right)=\sum_{i=1}^{n}\vec{F}_i\cdot\vec{G}_i \tag{4.1}$$

En efecto, para cada $i = 1,2,\ldots,m$:

$$\begin{aligned}\sum_{j=1}^{n} a_{ij}b_{ij} &= (a_{i1}b_{i1} + a_{i2}b_{i2} + \cdots + a_{in}b_{in})\\ &= (a_{i1}, a_{i2}, \ldots, a_{in}) \cdot (b_{i1}, b_{i2}, \ldots, b_{in})\\ &= \vec{F}_i \cdot \vec{G}_i.\end{aligned}$$

De donde se tiene (4.1). Sean ahora $A, B, C \in \mathfrak{M}_{m\times n}$ con sendas filas F_i, G_i y H_i, $i = 1,2,\ldots,m$; y $\lambda \in \mathbb{R}$. Entonces (note que utilizaremos las propiedades del producto punto en $\mathbb{R}^n$):

1.
$$\begin{aligned}\langle A, B\rangle &= \sum_{i=1}^{m} \vec{F}_i \cdot \vec{G}_i\\ &= \sum_{i=1}^{m} \vec{G}_i \cdot \vec{F}_i\\ &= \langle B, A\rangle.\end{aligned}$$

2.
$$\begin{aligned}\langle \lambda A, B\rangle &= \sum_{i=1}^{m} \left(\lambda \vec{F}_i\right) \cdot \vec{G}_i\\ &= \sum_{i=1}^{m} \lambda \left(\vec{F}_i \cdot \vec{G}_i\right)\\ &= \lambda \sum_{i=1}^{m} \vec{F}_i \cdot \vec{G}_i\\ &= \lambda \langle A, B\rangle.\end{aligned}$$

3.
$$\begin{aligned}\langle A, B+C\rangle &= \sum_{i=1}^{m} \vec{F}_i \cdot \left(\vec{G}_i + H_i\right)\\ &= \sum_{i=1}^{m} \left(\vec{F}_i \cdot \vec{G}_i + \vec{F}_i \cdot \vec{H}_i\right)\\ &= \sum_{i=1}^{m} \vec{F}_i \cdot \vec{G}_i + \sum_{i=1}^{m} \vec{F}_i \cdot \vec{H}_i\\ &= \langle A, B\rangle + \langle A, C\rangle.\end{aligned}$$

4.
$$\begin{aligned}\langle A, A\rangle &= \sum_{i=1}^{m} \vec{F}_i \cdot \vec{F}_i\\ &= \sum_{i=1}^{m} \|\vec{F}_i\|^2 \geq 0.\end{aligned}$$

5. Claramente, si $A = \mathscr{O}$, entonces, $\langle A, A\rangle = 0$. Supongamos que $\langle A, A\rangle = 0$, entonces,

$$0 = \langle A, A\rangle = \sum_{i=1}^{m} \|\vec{F}_i\|^2,$$

de donde $\|\vec{F}_i\|^2 = 0 \,\forall i = 1,2,\ldots,m$; luego, $a_{i1}^2 + a_{i2}^2 + \cdots a_{in}^2 = 0 \,\forall i = 1,2,\ldots,n$; por tanto, $a_{ij} = 0$ $\forall i, j$. Es decir, $A = \mathscr{O}$.

De 1, 2, 3, 4 y 5 $\langle A, B\rangle$ es un producto interior en $\mathfrak{M}_{m\times n}$.◀

Definición 4.2 ***(Traza de matrices cuadradas)*** *Sea $M \in \mathfrak{M}_{n\times n}$ una matriz cuadrada, se denota y define la traza de A como*

$$\operatorname{tra}(A) = \sum_{i=1}^{n} a_{ii}\,.$$

Es decir, la traza de A es la suma de los elementos de la diagonal de A.

► **Ejemplo 4.7** Si $A = \begin{bmatrix} -1 & 3 \\ 2 & 4 \end{bmatrix}$, entonces $\operatorname{tra}(A) = -1 + 4 = 3$.◄

► **Ejemplo 4.8** *(Producto interior en matrices mediante la traza).* Sean $A = [a_{ij}], B = [b_{ij}] \in \mathfrak{M}_{m\times n}$. Mostrar que el producto interior del ejemplo 4.6 definido por (4.1) se puede calcular por la fórmula

$$\langle A, B\rangle = \operatorname{tra}(B^t A)\,.◄$$

DEMOSTRACIÓN ■ Sea $C = B^t A$, con $C = [p_{kl}]$, y representemos por $\vec{F}_i$ y $\vec{G}_i$ las filas i de A y B, respectivamente, consideradas como vectores de $\mathbb{R}^n$. Entonces, dado que la componente p_{kl} de C es el producto de la fila k de B^t con la columna l de A y la fila k de B^t es la columna k de B, se tiene

$$p_{kl} = \begin{bmatrix} b_{1k} & b_{2k} & \cdots & b_{mk} \end{bmatrix} \begin{bmatrix} a_{1l} \\ a_{2l} \\ \vdots \\ a_{ml} \end{bmatrix}.$$

Entonces, los elementos de la diagonal de C son

$$p_{kk} = b_{1k}a_{1k} + b_{2k}a_{2k} + \cdots + b_{mk}a_{mk}.$$

Por tanto,

$$\begin{aligned}
\operatorname{tra}(C) &= \sum_{k=1}^{n} (b_{1k}a_{1k} + b_{2k}a_{2k} + \cdots + b_{mk}a_{mk}) \\
&= \sum_{k=1}^{n} b_{1k}a_{1k} + \sum_{k=1}^{n} b_{2k}a_{2k} + \cdots + \sum_{k=1}^{n} b_{mk}a_{mk} \\
&= \vec{G}_1 \cdot \vec{F}_1 + \vec{G}_2 \cdot \vec{F}_2 + \cdots + \vec{G}_m \cdot \vec{F}_m \\
&= \vec{F}_1 \cdot \vec{G}_1 + \vec{F}_2 \cdot \vec{G}_2 + \cdots + \vec{F}_m \cdot \vec{G}_m \\
&= \sum_{i=1}^{m} \vec{F}_i \cdot \vec{G}_i;
\end{aligned}$$

esto es,

$$\langle A, B\rangle = \sum_{i=1}^{m} \vec{F}_i \cdot \vec{G}_i = \operatorname{tra}(B^t A) \quad ■ \tag{4.2}$$

❍ **Nota 4.1** Observe que el producto interior en las matrices dado en (4.2) es una generalización del producto punto de vectores; ya que el producto punto de vectores, $\vec{u}\cdot\vec{v}$, se puede calcular mediante el producto matricial

$$(\vec{u})^t\,\vec{v}$$

al escribir los vectores $\vec{u}$ y $\vec{v}$ como matrices columna.

► **Ejemplo 4.9** *(Producto interior en un espacio de sucesiones).* Sea ℓ_2 el conjunto de las sucesiones de números reales (a_n) de cuadrado sumable; esto es,[4]

$$\ell_2 = \{(a_n)\,|\,\sum_{n=1}^{\infty} a_n^2 \text{ es una serie convergente}\}.$$

1. Probar que si $(a_n),(b_n)\in\ell_2$, entonces la serie

$$\sum_{n=1}^{\infty} a_n b_n$$

converge (absolutamente).

2. Mostrar que ℓ_2 es un subespacio vectorial del espacio de sucesiones reales[5] $\mathbb{R}^{\infty}$.
3. Sea, para cualquier par $(a_n),(b_n)\in\ell_2$,

$$\langle (a_n),(b_n)\rangle = \sum_{n=1}^{\infty} a_n b_n$$

(lo cual está bien definido por el inciso anterior). Demostrar que éste es un producto interior en ℓ_2.◄

DEMOSTRACIÓN ■ 1. Sean $(a_n),(b_n)$ un par de sucesiones en ℓ_2; entonces, $\sum_{n=1}^{\infty} a_n^2$ y $\sum_{n=1}^{\infty} b_n^2$ son series convergentes. De (3.5) del lema 3.1 (cfr. pág. 123) tenemos

$$2\,|a_n|\,|b_n| \le a_n^2 + b_n^2 \quad \forall n.$$

Por lo que la serie de términos no negativos $|a_n|\,|b_n|$ es convergente[6] y, por tanto, la serie $\sum_{i=1}^{\infty} a_n b_n$ es (absolutamente) convergente.[7]

2. *(i)* Claramente la sucesión constante cero ($a_n = 0\,\forall n$) pertenece a ℓ_2.

(ii) De la propiedad de la desigualdad triangular del valor absoluto,

$$|a_n + b_n| \le |a_n| + |b_n|,$$

se sigue que

$$\begin{aligned} |a_n+b_n|^2 &\le (|a_n|+|b_n|)^2 \\ &\le |a_n|^2 + 2\,|a_n|\,|b_n| + |b_n|^2 \end{aligned}$$

---

[4] Naturalmente se puede considerar, donde convenga, ℓ_2 el espacio de sucesiones (a_n) tales que $\sum_{n=0}^{\infty} a_n^2$ converge y $\langle (a_n),(b_n)\rangle = \sum_{n=0}^{\infty} a_n b_n$.

[5] Cfr. ejemplo 3.12, página 133.

[6] Recuerde que si $0 \le \alpha_n \le \beta_n$ $\forall n$, entonces, la convergencia de la serie $\sum_{n=1}^{\infty}\beta_n$ implica la convergencia de la serie $\sum_{n=1}^{\infty}\alpha_n$.

[7] Recuerde que la convergencia de la serie $\sum_{n=1}^{\infty}|\alpha_n|$ implica la convergencia de la serie $\sum_{n=1}^{\infty}\alpha_n$.

y, por ende, la serie

$$\sum_{n=1}^{\infty} |a_n + b_n|^2$$

converge. Ya que toda serie que converge absolutamente es convergente, se deduce que

$$\sum_{n=1}^{\infty} (a_n + b_n)^2$$

converge; es decir, $(a_n + b_n) \in \ell_2$.

(iii) Si $\alpha \in \mathbb{R}$ y $(a_n) \in \ell_2$, entonces $\sum_{n=1}^{\infty} a_n^2$ es convergente. Luego,

$$\begin{aligned} \sum_{n=1}^{\infty} (\lambda a_n)^2 &= \sum_{n=1}^{\infty} \lambda^2 a_n^2 \\ &= \lambda^2 \sum_{n=1}^{\infty} a_n^2 \end{aligned}$$

y, por tanto, $\lambda (a_n) \in \ell_2$.

De *(i)*, *(ii)* y *(iii)* se concluye que ℓ_2 es un subespacio del espacio vectorial de las sucesiones y por tanto un espacio vectorial.

3. Sean $(a_n), (b_n), (c_n) \in \ell_2$ y $\lambda \in \mathbb{R}$; entonces:

(a)
$$\begin{aligned} \langle (a_n), (b_n) \rangle &= \sum_{n=1}^{\infty} a_n b_n \\ &= \sum_{n=1}^{\infty} b_n a_n \\ &= \langle (b_n), (a_n) \rangle . \end{aligned}$$

(b)
$$\begin{aligned} \langle (\lambda a_n), (b_n) \rangle &= \sum_{n=1}^{\infty} (\lambda a_n) b_n \\ &= \lambda \sum_{n=1}^{\infty} a_n b_n \\ &= \lambda \langle (a_n), (b_n) \rangle . \end{aligned}$$

(c)
$$\begin{aligned} \langle (a_n), (b_n) + (c_n) \rangle &= \langle (a_n), (b_n + c_n) \rangle \\ &= \sum_{n=1}^{\infty} a_n (b_n + c_n) \\ &= \sum_{n=1}^{\infty} (a_n b_n + a_n c_n) \\ &= \sum_{n=1}^{\infty} a_n b_n + \sum_{n=1}^{\infty} a_n c_n \\ &= \langle (a_n), (b_n) \rangle + \langle (a_n), (c_n) \rangle . \end{aligned}$$

(d)
$$\langle (a_n), (a_n) \rangle = \sum_{n=1}^{\infty} a_n^2 \geq 0.$$

(e) Claramente, si (a_n) es la sucesión constante cero, $\langle (a_n), (a_n) \rangle = 0$. Supongamos inversamente que $\langle (a_n), (a_n) \rangle = 0$, para alguna sucesión $(a_n) \in \ell_2$, entonces,

$$a_m^2 \le \sum_{n=1}^{\infty} a_n^2 = 0;$$

de donde $a_m = 0\ \forall m$ y, por tanto, (a_n) es la sucesión constante cero. ■

▶ **Ejemplo 4.10** Sean las sucesiones $A = (1/n)$ y $B = (1/(n+1))$; i.e., las sucesiones $\{1, 1/2, 1/3, \ldots\}$ y $\{1/2, 1/3, 1/4, \ldots\}$.

1. Mostrar que $A, B \in \ell_2$.
2. Calcular

$$\langle (1/n), (1/(n+1)) \rangle . ◀$$

Solución 1. Sea $f(x) = 1/x^2$, $x \in [1, \infty)$. Entonces, f es decreciente en $[1, \infty)$, pues $f'(x) = -1/x^2 < 0$ en $[1, \infty)$; además,

$$\begin{aligned} \int_1^{\infty} f(x)dx &= \int_1^{\infty} \frac{dx}{x^2} \\ &= \lim_{r\to\infty} \int_1^r \frac{dx}{x^2} \\ &= \lim_{r\to\infty} \left[-\frac{1}{x} \right]_{x=1}^{x=r} \\ &= \lim_{r\to\infty} \left[-\frac{1}{r} + 1 \right] = 1. \end{aligned}$$

Por el criterio de la integral,[8] la serie

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \tag{4.3}$$

converge. Dado que la serie

$$\sum_{n=1}^{\infty} \frac{1}{(n+1)^2} \tag{4.4}$$

se obtiene de la serie (4.3) suprimiendo el primer término y la serie (4.3) es convergente, se deduce que la serie (4.4) es también convergente. Por tanto, las sucesiones A y B pertenecen al espacio ℓ_2.

2. Sea

$$\frac{1}{k(k+1)} = \frac{a}{k} + \frac{b}{k+1},$$

[8] Si f es una función no negativa y decreciente en el intervalo $[1, \infty)$ y la integral impropia $\int_1^{\infty} f(x)dx$ converge, entonces la serie $\sum_{n=1}^{\infty} f(n)$ converge.

entonces

$$1 = a(k+1) + bk = (a+b)k + a.$$

De donde

$$\begin{aligned} a &= 1, \\ a+b &= 0; \end{aligned}$$

por tanto, $b = -1$. Así que

$$\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}.$$

Entonces,

$$\begin{aligned} S_n &= \sum_{k=1}^{n} \frac{1}{k(k+1)} \\ &= \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1} \right) \\ &= 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \cdots - \frac{1}{n} + \frac{1}{n} - \frac{1}{n+1} \\ &= 1 - \frac{1}{n+1}. \end{aligned}$$

Luego,

$$\begin{aligned} \langle A, B \rangle &= \sum_{n=1}^{\infty} \frac{1}{n(n+1)} \\ &= \lim_{n\to\infty} S_n \\ &= \lim_{n\to\infty} \left(1 - \frac{1}{n+1} \right) \\ &= 1. \end{aligned}$$ ✔

❍ **Nota 4.2** También en el espacio ℓ_2 el producto interior, como en el caso del espacio $C[a,b]$, tiene un origen intuitivo en el producto punto de $\mathbb{R}^N$. En efecto, si (a_n) y (b_n) son un par de sucesiones en ℓ_2, sean $\vec{u}_N = (a_1, a_2, \ldots, a_N)$ y $\vec{v}_N = (b_1, b_2, \ldots, b_N)$, entonces

$$\begin{aligned} \langle (a_n), (b_n) \rangle &= \sum_{n=0}^{\infty} a_n b_n \\ &= \lim_{N\to\infty} \sum_{k=0}^{N} a_k b_k \\ &= \lim_{N\to\infty} \vec{u}_N \cdot \vec{v}_N. \end{aligned}$$

El lector puede verificar, sin mucha dificultad, las siguientes propiedades del producto interior, que son consecuencias inmediatas de la definición 4.1 y que hacemos patentes en el siguiente teorema.

Teorema 4.1 ***(Propiedades del producto interior)*** *Si* **E** *es un espacio vectorial y* $\langle \cdot , \cdot \rangle$ *es un producto interior en* **E**, *entonces* $\forall \vec{u}, \vec{v}, \vec{w} \in \mathbf{E}$ *y* $\forall \alpha, \beta \in \mathbb{R}$ *se cumple:*

1. $\langle \vec{u}, \lambda \vec{v} \rangle = \lambda \langle \vec{u}, \vec{v} \rangle$.
2. $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$.
3. $\langle \vec{u}, \alpha \vec{v} + \beta \vec{w} \rangle = \alpha \langle \vec{u}, \vec{v} \rangle + \beta \langle \vec{u}, \vec{w} \rangle$.
4. $\langle \alpha \vec{u} + \beta \vec{v}, \vec{w} \rangle = \alpha \langle \vec{u}, \vec{w} \rangle + \beta \langle \vec{u}, \vec{w} \rangle$.
5. $\langle \vec{u} - \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle - \langle \vec{v}, \vec{w} \rangle$.
6. $\langle \vec{u}, \vec{0}_{\mathbf{E}} \rangle = 0$, $\forall \vec{u} \in \mathbf{E}$.
7. $\langle \vec{u}, \vec{v} \rangle = 0 \; \forall \vec{v} \in \mathbf{E} \Rightarrow \vec{u} = \vec{0}_{\mathbf{E}}$.

4.1.2 Ortogonalidad y norma inducida por el producto interior

En este apartado extenderemos los conceptos de ortogonalidad (perpendicularidad) y norma a un espacio con producto escalar a partir del propio producto interior, como se hizo en el espacio $\mathbb{R}^n$. El concepto de ortogonalidad, definido en términos del producto interior nulo de los vectores involucrados, quedará justificado *a posteriori* al definir el ángulo entre vectores una vez que hayamos probado la desigualdad de Schwarz en estos espacios. A partir de aquí y hasta que terminemos esta sección, supondremos que **E** es un espacio con producto interior $\langle \cdot , \cdot \rangle$.

Definición 4.3 *(Ortogonalidad en espacios con producto interior) Se dice que* $\vec{u}, \vec{v} \in \mathbf{E}$ *son* **ortogonales** *(perpendiculares) si* $\langle \vec{u}, \vec{v} \rangle = 0$. *Si* $\vec{u}$ *y* $\vec{v}$ *son ortogonales, escribiremos* $\vec{u} \perp \vec{v}$.

► **Ejemplo 4.11** En $C[0, \pi]$, si $f(x) = \operatorname{sen}(x)$ y $g(x) = \cos(x)$,

$$\begin{aligned} \langle \operatorname{sen} x, \cos x \rangle &= \int_0^\pi \operatorname{sen} x \cos x \, dx \\ &= \left[\frac{\operatorname{sen}^2 x}{2} \right]_0^\pi \\ &= 0. \end{aligned}$$

Por tanto, $f \perp g$. Es decir, las funciones seno y coseno son ortogonales (perpendiculares) en $C[0, \pi]$.◄

► **Ejemplo 4.12** En $\mathfrak{M}_{2\times 2}$, con el producto interior definido en (4.2), si

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ -3 & 2 \end{bmatrix},$$

$$\langle A, B \rangle = (1)(1) + (1)(-1) + (2)(-3) + (3)(2) = 0$$

o, equivalentemente,

$$\begin{aligned}\langle A,B\rangle &= \operatorname{tra}(B^t A)\\ &= \operatorname{tra}\left(\begin{bmatrix} 1 & -3\\ -1 & 2\end{bmatrix}\begin{bmatrix} 1 & 1\\ 2 & 3\end{bmatrix}\right)\\ &= \operatorname{tra}\left(\begin{bmatrix} -5 & -8\\ 3 & 5\end{bmatrix}\right)\\ &= 0.\end{aligned}$$

Por lo que $A \perp B$ en este espacio.◄

Las siguientes propiedades de ortogonalidad, enunciadas en el teorema 4.2, son sencillas de demostrar si se utilizan las propiedades del producto interior. Se deja al lector su demostración como ejercicio.

Teorema 4.2 ***(Propiedades de ortogonalidad)*** *Si $\vec{u},\vec{v},\vec{w} \in \mathbf{E}$, y $\alpha,\beta \in \mathbb{R}$, entonces,*

1. $\vec{u} \perp \vec{v}, \vec{u} \perp \vec{w} \Rightarrow \vec{u} \perp \alpha\vec{u} + \beta\vec{v}$.
2. *Si $S = \operatorname{gn}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m)$ es un subespacio de $\mathbf{E}$ y $\vec{u} \perp \vec{v}_i$ para cada $i = 1, 2, \ldots, m$, entonces $\vec{u} \perp \vec{v}$ $\forall \vec{v} \in S$.*
3. $\vec{u} \perp \vec{v}\ \ \forall \vec{v} \in \mathbf{E} \Leftrightarrow \vec{u} = \vec{0}_{\mathbf{E}}$ *(el único vector en un espacio que es ortogonal a todos los demás es el vector cero, el neutro aditivo).*

Definición 4.4 ***(Norma inducida por un producto interior)*** *Se define la norma de un vector $\vec{u}$ en $\mathbf{E}$ como*

$$\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u}\rangle}\,.$$

A dicho número se le llama la ***norma inducida*** *por el producto interior $\langle \cdot, \cdot \rangle$ de* $\mathbf{E}$.

► **Ejemplo 4.13** Sea $A \in \mathfrak{M}_{m\times n}$, $A = [a_{ij}]$, con filas $\vec{F}_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, entonces,

$$\begin{aligned}\|A\|^2 &= \langle A, A\rangle = \sum_{i=1}^{m} \vec{F}_i \cdot \vec{F}_i\\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\,.\end{aligned}$$

Es decir,

$$\|A\| = \sqrt{\sum a_{ij}^2}\,,$$

donde la suma se efectúa sobre todas las componentes de la matriz A.◄

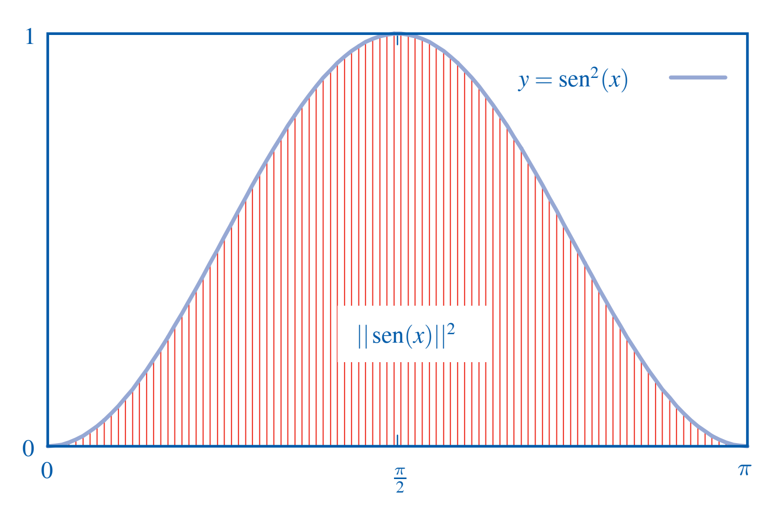


Figura 4-1 • Gráfica de la función $y = \text{sen}^2(x)$ y del área bajo esta curva en el intervalo $[0, \pi]$. La raíz cuadrada del valor de esta área es la norma de la función $y = \text{sen}(x)$ en el espacio $C[a, b]$.

▶ **Ejemplo 4.14** En $C[0, \pi]$

$$\begin{aligned} \|\,\text{sen}\,x\| &= \sqrt{\langle \text{sen}\,x, \text{sen}\,x\rangle} = \sqrt{\int_0^\pi \text{sen}^2 x dx} \\ &= \left[\int_0^\pi \frac{1 - \cos 2x}{2} dx\right]^{1/2} \\ &= \left[\frac{1}{2}x\Big|_0^\pi - \frac{1}{4}\,\text{sen}\,2x\Big|_0^\pi\right]^{1/2} = \sqrt{\frac{\pi}{2}}. \end{aligned}$$ ◀

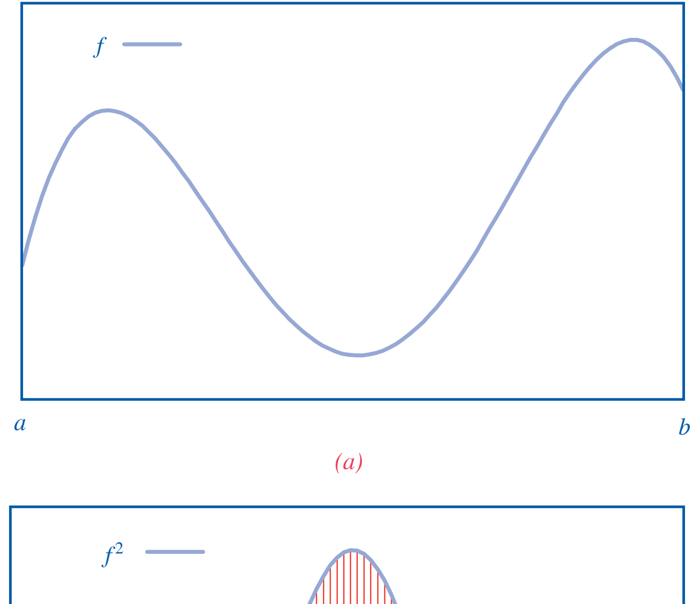
La figura 4.1 contiene la gráfica de la función $y = \text{sen}^2(x)$ y el área bajo esta curva en el intervalo $[0, \pi]$. La raíz cuadrada de esta área es la norma de la función $y = \text{sen}(x)$ en el espacio $C[0, \pi]$.

Ésta es la interpretación geométrica que tiene la norma inducida por el producto interior $\langle f, g\rangle = \int_a^b f(x)g(x)\,dx$ en el espacio de funciones continuas en $[a, b]$. La norma de una función f es la raíz cuadrada del área bajo la curva $y = f^2(x)$ en el intervalo $[a, b]$, como hacemos patente en la figura 4.2. Más adelante, cuando veamos el concepto de distancia entre vectores, daremos una interpretación más profunda de la norma inducida por el producto interior[9] (cfr. ejemplo 4.21 y la discusión ulterior en la pág. 258) que, *grosso modo*, establece que la norma de una función f en $C[a, b]$ estima el promedio de los valores cuadráticos que toma una función en el intervalo $[a, b]$.

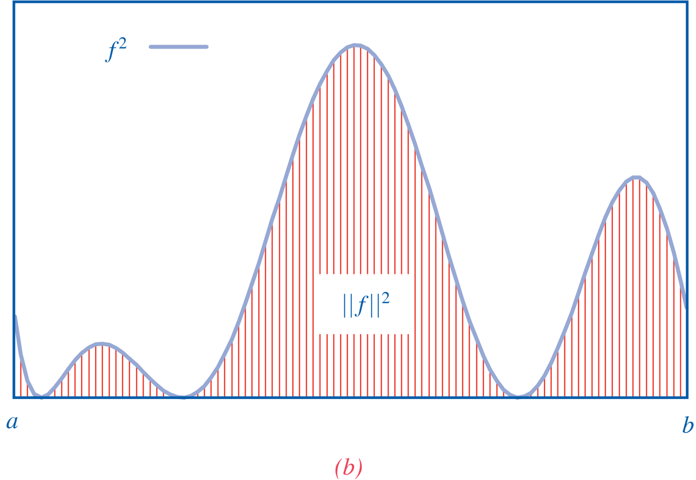
▶ **Ejemplo 4.15** En el espacio de sucesiones, sea $A = \left(\frac{1}{2^n}\right)$.

1. Mostrar que $A \in \ell_2$, el espacio de sucesiones de cuadrado sumable (cfr. ejemplo 4.9).
2. Calcular $\|A\|$, la norma inducida por el producto interior en ℓ_2. ◀

[9] Con frecuencia a la norma inducida por el producto interior en $C[a, b]$, $\left(\int_a^b (f(x))^2\,dx\right)^{1/2}$, se le llama **norma cuadrado medio** o **norma de promedio cuadrático**.



(a)



(b)

Figura 4-2 • *(a)* Gráfica de una función f. *(b)* Gráfica de la función f^2 y del área bajo esta curva en el intervalo $[a,b]$, cuyo valor es el cuadrado de la norma de f en el espacio $C[a,b]$.

Solución 1. Sea, para cada $n = 0, 1, 2, \ldots,$

$$\sigma_n = 1 + \frac{1}{2^2} + \frac{1}{2^4} + \cdots + \frac{1}{2^{2n-2}} + \frac{1}{2^{2n}}\ .$$

Entonces,

$$\frac{1}{2^2}\sigma_n = \frac{1}{2^2} + \frac{1}{2^4} + \cdots + \frac{1}{2^{2n}} + \frac{1}{2^{2n+2}}\ ;$$

por tanto,

$$\sigma_n - \frac{1}{2^2}\sigma_n = 1 - \frac{1}{2^{2n+2}}\ .$$

En consecuencia,

$$\sigma_n = \frac{4}{3}\left(1 - \frac{1}{2^{2n+2}}\right);$$

luego,

$$\begin{aligned}\lim_{n\to\infty}\sigma_n &= \lim_{n\to\infty}\frac{4}{3}\left(1-\frac{1}{2^{2n+2}}\right)\\ &= \frac{4}{3}\,.\end{aligned}$$

Por lo que,

$$\sum_{n=0}^{\infty}\left(\frac{1}{2^n}\right)^2=\frac{4}{3} \tag{4.5}$$

Esto es, $A\in\ell_2$.

2. De (4.5), se tiene

$$\langle A,A\rangle=\sum_{n=0}^{\infty}\left(\frac{1}{2^n}\right)^2=\frac{4}{3}$$

Por ende,

$$\|A\|=\sqrt{\langle A,A\rangle}=\sqrt{\frac{4}{3}}=\frac{2}{\sqrt{3}}=\frac{2}{3}\sqrt{3}.\quad ✔$$

Nuevamente tenemos, como consecuencia del concepto geométrico de ortogonalidad, el célebre teorema de Pitágoras en una versión mucho más general.

Teorema 4.3 ***(Teorema de Pitágoras en espacios con producto interior)*** *Si* **E** *es un espacio con producto interior* $\langle\cdot,\cdot\rangle$ *y* $\vec{x}\perp\vec{y}$ *en* **E**, *entonces,*

$$\|\vec{x}+\vec{y}\|^2=\|\vec{x}\|^2+\|\vec{y}\|^2.$$

DEMOSTRACIÓN ■ Como $\vec{x}\perp\vec{y}$, $\langle\vec{x},\vec{y}\rangle=0$. Entonces,

$$\begin{aligned}\|\vec{x}+\vec{y}\|^2 &= \langle\vec{x}+\vec{y},\vec{x}+\vec{y}\rangle\\ &= \langle\vec{x}+\vec{y},\vec{x}\rangle+\langle\vec{x}+\vec{y},\vec{y}\rangle\\ &= \langle\vec{x},\vec{x}\rangle+\langle\vec{y},\vec{x}\rangle+\langle\vec{x},\vec{y}\rangle+\langle\vec{y},\vec{y}\rangle\\ &= \|\vec{x}\|^2+2\langle\vec{x},\vec{y}\rangle+\|\vec{y}\|^2\\ &= \|\vec{x}\|^2+\|\vec{y}\|^2.\quad ■\end{aligned}$$

Es fácil ver, aplicando el teorema anterior, que si $\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_k\in\mathbf{E}$ son ortogonales entre sí, entonces,

$$\|\vec{x}_1+\vec{x}_2+\cdots+\vec{x}_k\|^2=\|\vec{x}_1\|^2+\|\vec{x}_2\|^2+\cdots+\|\vec{x}_k\|^2,$$

la cual es una versión más general del teorema anterior.

4.1.3 Desigualdad de Schwarz y ángulo entre vectores

Vector proyección

Si $\vec{u}$ y $\vec{v}$ son dos vectores en $\mathbb{R}^2$, el vector proyección de $\vec{u}$ sobre $\vec{v}$ se obtiene trazando una línea perpendicular a $\vec{v}$ que pase por $\vec{u}$, como se ilustra en la figura 4.3. Generalizamos este concepto a espacios con producto interior en la siguiente definición.

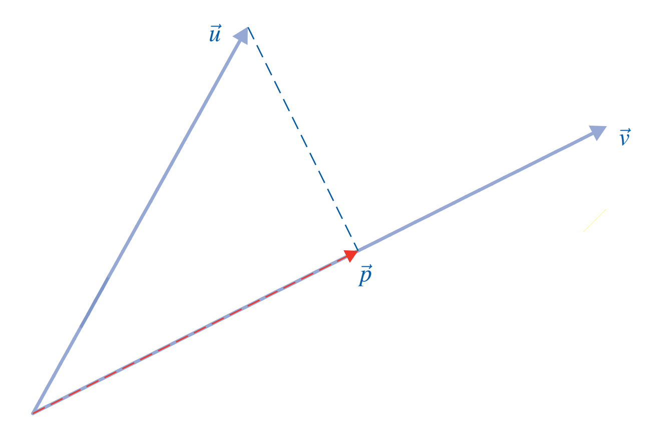


Figura 4-3 • El vector proyección $\vec{p}$ de un vector $\vec{u}$ sobre un vector $\vec{v}$ en $\mathbb{R}^2$.

Definición 4.5 *En un espacio con producto interior se define el vector proyección, $\vec{p}$, de $\vec{u}$ sobre $\vec{v}$ ($\vec{v} \neq \vec{0}_{\mathbf{E}}$), como aquel que satisface:*

1. $\vec{p} \in \text{gn}(\vec{v})$.
2. $\vec{u} - \vec{p} \perp \vec{v}$.

Por la primera condición, se tiene que $\vec{p} = \lambda\vec{v}$; y, por la segunda,

$$\begin{aligned} 0 &= \langle \vec{u} - \vec{p}, \vec{v} \rangle \\ &= \langle \vec{u}, \vec{v} \rangle - \langle \vec{p}, \vec{v} \rangle \\ &= \langle \vec{u}, \vec{v} \rangle - \lambda \langle \vec{v}, \vec{v} \rangle \\ &= \langle \vec{u}, \vec{v} \rangle - \lambda \|\vec{v}\|^2. \end{aligned}$$

Lo que implica

$$\lambda = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|^2},$$

y, entonces,

$$\vec{p} = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|^2}\,\vec{v}.$$

Con lo que hemos probado el siguiente teorema.

Teorema 4.4 *Si $\vec{v}, \vec{u} \in \mathbf{E}$ (con $\vec{v} \neq \vec{0}_{\mathbf{E}}$), entonces el vector proyección $\vec{p}$ de $\vec{u}$ sobre $\vec{v}$ está dado por*

$$\vec{p} = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|^2} \vec{v} \tag{4.6}$$

▶ **Ejemplo 4.16** Sean $\vec{u} = (1,1)$ y $\vec{v} = (2,0)$, entonces,

$$\begin{aligned} \vec{p} &= \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2} \vec{v} \\ &= \frac{(1,1) \cdot (2,0)}{\|(2,0)\|^2} (2,0) \\ &= \frac{1}{2} (2,0) \\ &= (1,0). \end{aligned}$$ ◀

▶ **Ejemplo 4.17** Si $f(x) = e^x$ y $g(x) = e^{-2x}$ en $C[0,1]$, con el producto interior definido en el ejemplo 4.3,[10] entonces, la proyección de f sobre g está dada por

$$p(x) = \frac{\langle f(x), g(x) \rangle}{\|g(x)\|^2} g(x).$$

Por una parte, tenemos

$$\begin{aligned} \langle f(x), g(x) \rangle &= \int_0^1 e^x e^{-2x} dx \\ &= \int_0^1 e^{-x} dx \\ &= 1 - e^{-1} \end{aligned}$$

y, por otro lado,

$$\begin{aligned} \|g(x)\|^2 &= \langle g(x), g(x) \rangle \\ &= \int_0^1 \left(e^{-2x}\right)^2 dx \\ &= \int_0^1 e^{-4x} dx \\ &= \frac{1}{4} \left(1 - e^{-4}\right). \end{aligned}$$

Luego,

$$p(x) = \frac{4\left(1 - e^{-1}\right)}{1 - e^{-4}} e^{-2x}.$$ ◀

[10] De aquí en adelante, en todos los ejemplos con los que se trabaje, se supondrá que los productos escalares en cada espacio son los que hemos definido en los ejemplos previos en este capítulo, a menos que se indique otra cosa.

▶ **Ejemplo 4.18** Si $A = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & -1 \\ 2 & 0 \end{bmatrix}$ en $\mathfrak{M}_{2\times 2}$, entonces

$$\langle A, B\rangle = 1$$

y

$$\|B\|^2 = 5$$

así que la proyección de A sobre B es

$$\begin{aligned} P &= \frac{\langle A,B\rangle}{\|B\|^2}B \\ &= \frac{1}{5}\begin{bmatrix} 0 & -1 \\ 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -\frac{1}{5} \\ \frac{2}{5} & 0 \end{bmatrix}. \end{aligned}$$ ◀

▶ **Ejemplo 4.19** Sean $A = \left(\frac{1}{3^n}\right)$ y $B = \left(\frac{1}{2^n}\right)$, puesto que

$$\begin{aligned} \left(\frac{1}{2^n}\right)^2 &\le \frac{1}{2^n} \quad \text{y} \\ \left(\frac{1}{3^n}\right)^2 &\le \frac{1}{3^n} \end{aligned}$$

para todo n y las series geométricas $\sum_{n=0}^{\infty}\frac{1}{2^n}$ y $\sum_{n=0}^{\infty}\frac{1}{3^n}$ son convergentes, se sigue que $\sum_{n=0}^{\infty}\left(\frac{1}{2^n}\right)^2$ y $\sum_{n=0}^{\infty}\left(\frac{1}{3^n}\right)^2$ son convergentes; por tanto $A, B \in \ell_2$. Hallar la proyección de A sobre B. ◀

Solución Tenemos que[11]

$$\begin{aligned} \langle A,B\rangle &= \sum_{n=0}^{\infty}\frac{1}{3^n}\frac{1}{2^n} \\ &= \sum_{n=0}^{\infty}\frac{1}{(2\cdot 3)^n} \\ &= \sum_{n=0}^{\infty}\left(\frac{1}{6}\right)^n \\ &= \frac{1}{1-\frac{1}{6}} \\ &= \frac{6}{5} \end{aligned}$$

[11] Aquí hemos utilizado el conocido hecho de que la serie geométrica $\sum_{n=0}^{\infty} r^n$ converge si y sólo si $|r| < 1$; y que en tal caso $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$. Pudimos haber utilizado en los ejemplos precedentes esta propiedad de las series geométricas, pero hemos empleado otros criterios más básicos de series con el objetivo de que el lector pueda repasarlos.

y, por el ejemplo 4.15, sabemos que

$$\|B\|^2 = \frac{4}{3}.$$

Entonces,

$$\begin{aligned} P &= \frac{\langle A,B\rangle}{\|B\|^2}B \\ &= \frac{6/5}{4/3}B; \end{aligned}$$

esto es,

$$P = \left(\frac{9}{10}\frac{1}{2^n}\right)$$

es la sucesión proyección de A sobre B. ✔

Como vimos en el capítulo anterior, la desigualdad que permite definir ángulos entre vectores de $\mathbb{R}^n$ es la de Schwarz. Corresponde establecer ahora su versión en espacios vectoriales con producto escalar. Antes de probar la desigualdad de Schwarz, observemos que, para cualquier $\vec{u} \in \mathbf{E}$ y para todo $\lambda \in \mathbb{R}$,

$$\begin{aligned} \|\lambda\vec{u}\|^2 &= \langle \lambda\vec{u}, \lambda\vec{u}\rangle \\ &= \lambda\langle \vec{u}, \lambda\vec{u}\rangle \\ &= \lambda^2\langle \vec{u}, \vec{u}\rangle \\ &= \lambda^2\|\vec{u}\|^2; \end{aligned}$$

es decir,

$$\|\lambda\vec{u}\| = |\lambda|\|\vec{u}\|.$$

Resultado que emplearemos en la demostración de esta desigualdad.

Desigualdad de Schwarz

Teorema 4.5 *Si* $\mathbf{E}$ *es un espacio con producto interior* $\langle\cdot,\cdot\rangle$, *entonces,*

$$|\langle \vec{u}, \vec{v}\rangle| \le \|\vec{u}\|\,\|\vec{v}\| \tag{4.7}$$

$\forall \vec{u}, \vec{v} \in \mathbf{E}$.

DEMOSTRACIÓN ■ Si $\vec{u} = \vec{0}_{\mathbf{E}}$ o $\vec{v} = \vec{0}_{\mathbf{E}}$, claramente (4.7) es cierta. Supongamos que ninguno de los dos vectores es nulo. Sea

$$\vec{p} = \frac{\langle \vec{u}, \vec{v}\rangle}{\|\vec{v}\|^2}\vec{v}$$

el vector proyección de $\vec{u}$ sobre $\vec{v}$. Entonces $\vec{u}-\vec{p}\perp\vec{p}$. Por el teorema de Pitágoras:

$$\|\vec{u}\|^2 = \|(\vec{u}-\vec{p})+\vec{p}\|^2 = \|\vec{u}-\vec{p}\|^2+\|\vec{p}\|^2.$$

Lo cual implica

$$\|\vec{u}\|^2 = \|\vec{u}-\vec{p}\|^2+\|p\|^2 \geq \|p\|^2;$$

es decir,

$$\begin{aligned} \|\vec{u}\|^2 &\geq \|\vec{p}\|^2 \\ &= \left\| \frac{\langle\vec{u},\vec{v}\rangle}{\|v\|^2}\vec{v} \right\|^2 \\ &= \frac{|\langle\vec{u},\vec{v}\rangle|^2}{\|v\|^4}\|\vec{v}\|^2 \\ &= \frac{|\langle\vec{u},\vec{v}\rangle|^2}{\|v\|^2}. \end{aligned}$$

De donde,

$$\|\vec{u}\|^2\|\vec{v}\|^2 \geq |\langle\vec{u},\vec{v}\rangle|^2$$

y, por tanto,

$$\|\vec{u}\|\|\vec{v}\| \geq |\langle\vec{u}|\vec{v}\rangle|. \quad \blacksquare$$

❍ **Nota 4.3** Se puede probar que hay igualdad en (4.7) si y sólo si $\vec{u}$ y $\vec{v}$ son linealmente dependientes (L.D.).

Propiedades de la norma inducida por el producto interior

En la subsección 3.1.4, vimos que la norma inducida por el producto punto en $\mathbb{R}^n$, esto es $\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}}$, tiene las propiedades enunciadas en el teorema 3.3 (cfr. pág. 127), las cuales se deducen a partir de las propiedades del producto punto en $\mathbb{R}^n$, que son en las que nos basamos para definir un producto interior en general. Así que, como es de esperar, la norma inducida por un producto interior, es decir, $\|\vec{u}\| = \sqrt{\langle\vec{u},\vec{u}\rangle}$, también debe tener las mismas propiedades que su caso particular en $\mathbb{R}^n$. En el siguiente teorema hacemos patentes las características esenciales de la norma inducida por un producto interior; a partir de ellas se puede deducir cualquier otra propiedad de dicha norma.

Teorema 4.6 *Sea* **E** *un espacio vectorial con producto interior* $\langle\cdot,\cdot\rangle$. *Entonces, si* $\|\vec{u}\| = \sqrt{\langle\vec{u},\vec{u}\rangle}$ *es la norma inducida por el producto interior, se tiene:*

1. $\|\vec{u}\| \geq 0 \quad \forall\vec{u}\in\mathbf{E}$.
2. $\|\vec{u}\| = 0 \Leftrightarrow \vec{u} = \vec{0}_{\mathbf{E}}$.
3. $\|\lambda\vec{u}\| = |\lambda|\,\|\vec{u}\| \quad \forall\vec{u}\in\mathbf{E}, \forall\lambda\in\mathbb{R}$.
4. $\|\vec{u}+\vec{v}\| \leq \|\vec{u}\|+\|\vec{v}\| \quad \forall\vec{u},\vec{v}\in\mathbf{E}$ (**Desigualdad triangular**).

DEMOSTRACIÓN ■ 1. $\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}} \geq 0\ \forall \vec{u} \in \mathbf{E}$.

2. $\|\vec{0}_{\mathbf{E}}\| = \sqrt{\langle \vec{0}_{\mathbf{E}}, \vec{0}_{\mathbf{E}} \rangle} = \sqrt{0} = 0$. Supongamos que $\|\vec{u}\| = 0$, entonces $\langle \vec{u}, \vec{u} \rangle = 0$ y, por tanto, $\vec{u} = \vec{0}_{\mathbf{E}}$.

3. Si $\lambda \in \mathbb{R}$ y $\vec{u} \in \mathbf{E}$, entonces,

$$\begin{aligned} \|\lambda\vec{u}\|^2 &= \langle \lambda\vec{u}, \lambda\vec{u} \rangle \\ &= \lambda \langle \vec{u}, \lambda\vec{u} \rangle \\ &= \lambda^2 \langle \vec{u}, \vec{u} \rangle \\ &= \lambda^2 \|\vec{u}\|^2, \end{aligned}$$

de donde,

$$\|\lambda\vec{u}\| = |\lambda|\,\|\vec{u}\|\,.$$

4. Sean $\vec{u}, \vec{v} \in \mathbf{E}$, entonces,

$$\begin{aligned} \|\vec{u}+\vec{v}\|^2 &= \langle \vec{u}+\vec{v}, \vec{u}+\vec{v} \rangle \\ &= \langle \vec{u}+\vec{v}, \vec{u} \rangle + \langle \vec{u}+\vec{v}, \vec{v} \rangle \\ &= \langle \vec{u}, \vec{u} \rangle + \langle \vec{v}, \vec{u} \rangle + \langle \vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{v} \rangle \\ &= \|\vec{u}\|^2 + 2\langle \vec{u}, v \rangle + \|\vec{v}\|^2\,. \end{aligned}$$

De la desigualdad de Schwarz (4.7)

$$\langle \vec{u}, v \rangle \leq \|\vec{u}\|\,\|\vec{v}\|\,.$$

Así que,

$$\begin{aligned} \|\vec{u}+\vec{v}\|^2 &= \|\vec{u}\|^2 + 2\langle \vec{u}, v \rangle + \|\vec{v}\|^2 \\ &\leq \|\vec{u}\|^2 + 2\|\vec{u}\|\,\|\vec{v}\| + \|\vec{v}\|^2 \\ &= (\|\vec{u}\| + \|\vec{v}\|)^2\,. \end{aligned}$$

Luego,

$$\|\vec{u}+\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\| \tag{4.8}$$

■

❍ **Nota 4.4** Se puede probar que hay igualdad en (4.8) si y sólo si uno de los vectores es múltiplo escalar no negativo del otro.

Distancia en espacios con producto interior

Una vez que se tiene definido un producto interior en un espacio vectorial, se puede introducir el concepto de distancia entre vectores a través de la norma inducida por el producto interior; como hacemos patente en la siguiente definición.

Definición 4.6 *Si $\vec{u}$ y $\vec{v}$ son un par de vectores en el espacio* **E**, *se define y denota la* ***distancia*** *entre ellos como*

$$\mathrm{d}(\vec{u},\vec{v}) = \|\vec{u}-\vec{v}\|.$$

▶ **Ejemplo 4.20** En $\mathfrak{M}_{2\times 2}$ si $A = \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, entonces,

$$\begin{aligned} \mathrm{d}(A,B) &= \|A-B\| \\ &= \left\| \begin{bmatrix} -3 & 0 \\ 1 & -1 \end{bmatrix} \right\| \\ &= \sqrt{(-3)^2+(1)^2+(-1)^2} \\ &= \sqrt{11}. \end{aligned}$$ ◀

▶ **Ejemplo 4.21** Sean f y g un par de funciones continuas en $[a,b]$. Entonces la distancia entre ellas, medida con la norma inducida por el producto interior, es

$$\mathrm{d}(f,g) = \|f-g\| = \left(\int_a^b (f(x)-g(x))^2\,dx\right)^{1/2}.$$

Por ejemplo, si $f(x)=1$ y $g(x)=x$, entonces en $C[0,1]$

$$\begin{aligned} \mathrm{d}(f,g) &= \left(\int_0^1 (1-x)^2\,dx\right)^{1/2} = \left(-\frac{(1-x)^3}{3}\Bigg|_{x=0}^{x=1}\right)^{1/2} \\ &= \frac{1}{\sqrt{3}} \end{aligned}$$ ◀

La distancia entre dos funciones en $C[a,b]$ es la raíz cuadrada del área de la diferencia cuadrática entre ellas en el intervalo $[a,b]$, como se ilustra en la figura 4-4.

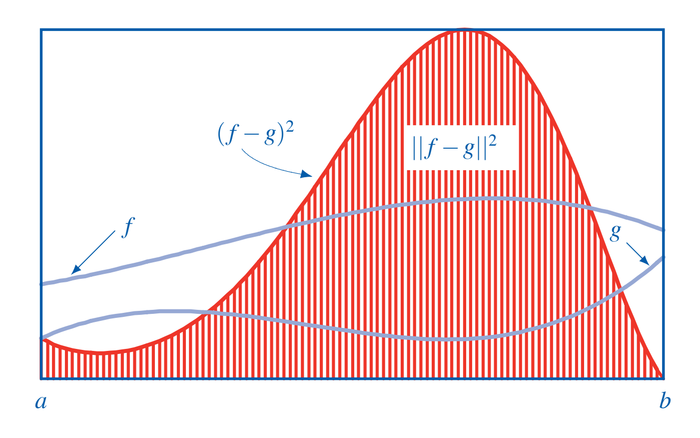


Figura 4-4 • Gráficas de las funciones f, g, $(f-g)^2$ y del área bajo esta curva en el intervalo $[a,b]$.

En realidad, la distancia entre dos funciones en $C[a,b]$ es una medida de la variación promedio cuadrática entre ellas. Para entender esto, primero supongamos que $\varphi \in C[a,b]$; y sean N un entero no negativo; $h = (b-a)/N$; $x_k = a + kh$, $k = 0,1,\ldots,N$; y $x_k^* \in [x_{k-1}, x_k]$ un punto de este k-ésimo subintervalo, $k = 1,2,\ldots,N$. Sabemos de la definición de integral que

$$\begin{aligned}\int_a^b \varphi(x)dx &= \lim_{N\to\infty} \sum_{k=1}^{N} \varphi(x_k^*)h \\ &= (b-a) \lim_{N\to\infty} \frac{\sum_{k=1}^{N} \varphi(x_k^*)}{N}\end{aligned}$$

Por lo que,

$$\lim_{N\to\infty} \frac{\sum_{k=1}^{N} \varphi(x_k^*)}{N} = \frac{1}{b-a}\int_a^b \varphi(x)dx\,.$$

De aquí que, para N grande,

$$\frac{\sum_{k=1}^{N} \varphi(x_k^*)}{N} \approx \frac{1}{b-a}\int_a^b \varphi(x)dx \tag{4.9}$$

Luego, el promedio de los valores de la función en los puntos x_k^* es aproximadamente el área bajo la curva $y = \varphi(x)$ en el intervalo $[a,b]$, dividida entre la longitud de este intervalo; y esta aproximación es más precisa a medida que es mayor el número de los puntos x_k^*. Para ilustrar este hecho numéricamente, consideremos el caso particular de la función $\varphi(x) = x^2$ en el intervalo $[0,1]$. Los valores x_k^*, que incluimos en la siguiente matriz, se eligieron aleatoriamente del intervalo $[0,1]$ para $N = 9$:

$$\begin{matrix} 0.3365 & 0.2897 & 0.9254 \\ 0.2481 & 0.1814 & 0.3306 \\ 0.4026 & 0.5126 & 0.8847 \end{matrix}$$

Entonces,

$$\frac{1}{9}\sum_{k=1}^{9} \varphi(x_k^*) = 0.2662$$

N	100	400	900	1 600	2 500
$\frac{\sum_{k=1}^{N} \varphi(x_k^*)}{N}$	.3596	.3215	.3370	.3291	.3312
N	3 600	4 900	6 400	8 100	14 400
$\frac{\sum_{k=1}^{N} \varphi(x_k^*)}{N}$	.3427	.3301	.3362	.3390	.3327

Tabla 4-1 • Valores promedio de los $\varphi(x_k^*)$ para diversos valores de N con los puntos x_k^* elegidos aleatoriamente en el intervalo $[0,1]$ para cada N.

La tabla 4.1 incluye los promedios dados por el lado izquierdo de (4.9) para distintos valores de N y para puntos x_k^* ,$k = 1,2,\ldots,N$, elegidos aleatoriamente en el intervalo $[0,1]$ para la función $\varphi(x) = x^2$ (los valores de los x_k^* no se incluyen por razones obvias de espacio).

El valor del lado derecho de (4.9) en este caso es

$$\int_0^1 x^2 dx = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}.$$

De la tabla 4.1 se puede ver que la tendencia es precisamente a $1/3$ cuando $N \to \infty$. Entonces, la interpretación que tiene $\|f-g\|$ en $C[a,b]$ va más allá que simplemente ser la raíz cuadrada del valor del área bajo la función $(f-g)^2$ en el intervalo $[a,b]$; sino que el cuadrado de esta norma dividido entre la longitud del intervalo es el límite del promedio de las diferencias al cuadrado entre estas funciones. Esto es,

$$\frac{\|f-g\|^2}{b-a} \approx \frac{\sum_{k=1}^{N}\left(f(x_k^*)-g(x_k^*)\right)^2}{N}$$

para N grande. Es por esto que, aunque $\|f-g\|$ no mide una diferencia puntal entre estas funciones, sí mide el error promedio cuadrático entre ellas; este enfoque es mucho más valioso para una gran variedad de aplicaciones, por lo que esta norma se emplea con gran frecuencia cuando se estudian espacios de funciones continuas.

Ángulos entre vectores en espacios con producto interior

De la desigualdad de Schwarz (4.7) se deduce que

$$-\|\vec{u}\|\|\vec{v}\| \leq \langle\vec{u},\vec{v}\rangle \leq \|\vec{u}\|\|\vec{v}\|$$

lo cual implica, si ambos vectores son no nulos,

$$-1 \leq \frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{u}\|\|\vec{v}\|} \leq 1.$$

Por tanto, es posible definir ángulo entre vectores en un espacio con producto interior a partir de la función arc cos, como lo hicimos en el caso $\mathbb{R}^2$ y, en general, en $\mathbb{R}^n$. Hacemos patente este concepto en la siguiente definición.

Definición 4.7 *Si* **E** *es un espacio con producto interior y $\vec{u},\vec{v} \in \mathbf{E}-\{\vec{0}_{\mathbf{E}}\}$, se define el ángulo entre $\vec{u}$ y $\vec{v}$ como:*

$$\theta = \arccos\left(\frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{u}\|\|\vec{v}\|}\right).$$

► **Ejemplo 4.22** En $C[0,1]$, si $f(x)=x^2$, $g(x)=x$,

$$\langle f,g\rangle = \int_0^1 f(x)g(x)dx = \int_0^1 x^2\cdot x dx = x^4/4\Big|_0^1 = 1/4;$$

$$\|f\| = \sqrt{\int_0^1 x^4 dx} = 1/\sqrt{5};$$

$$\|g\| = \sqrt{\int_0^1 x^2 dx} = \left[\sqrt{x^3/3} =\right]_0^1 = 1/\sqrt{3}.$$

Entonces,

$$\theta = \arccos\left(\frac{\langle f,g\rangle}{\|f\|\,\|g\|}\right) = \arccos\left(\frac{1/4}{\frac{1}{\sqrt{5}\sqrt{3}}}\right),$$

y, finalmente,

$$\theta = \arccos\left(\frac{\sqrt{15}}{4}\right) \approx 0.2526\,\text{rad} \approx 14.47°. ◀$$

▶ **Ejemplo 4.23** En $\mathfrak{M}_{2\times 2}$, sean $A = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$, entonces el ángulo entre estas matrices es

$$\begin{aligned} \theta &= \arccos\left(\frac{\langle A,B\rangle}{\|A\|\,\|B\|}\right) \\ &= \arccos\left(\frac{1}{\sqrt{2}\sqrt{2}}\right) \\ &= \arccos\left(\frac{1}{2}\right) \\ &= 60°. ◀ \end{aligned}$$

▶ **Ejemplo 4.24** En ℓ_2 sean las sucesiones[12] $A = \left(\left(\frac{\sqrt{2}}{2}\right)^n\right)$ y $B = \left(\left(\frac{\sqrt{2}}{10}\right)^n\right)$, entonces el ángulo entre estas sucesiones es

$$\theta = \arccos\left(\frac{\langle A,B\rangle}{\|A\|\,\|B\|}\right).$$

Tenemos que

$$\begin{aligned} \langle A,B\rangle &= \sum_{n=0}^{\infty}\left(\frac{\sqrt{2}}{2}\right)^n\left(\frac{\sqrt{2}}{10}\right)^n \\ &= \sum_{n=0}^{\infty}\left(\frac{\sqrt{2}\sqrt{2}}{2\cdot 10}\right)^n \\ &= \sum_{n=0}^{\infty}\left(\frac{1}{10}\right)^n \\ &= \frac{1}{1-\frac{1}{10}} \\ &= \frac{10}{9} \end{aligned}$$

[12] $((\sqrt{2}/2)^n)^2 = (1/2)^n$ y $((\sqrt{2}/10)^n)^2 = (1/50)^n$; puesto que $1/2 < 1$ y $1/50 < 1$, las series geométricas $\sum_{n=1}^{\infty}((\sqrt{2}/2)^n)^2$, $\sum_{n=1}^{\infty}((\sqrt{2}/10)^n)^2$ convergen y, por tanto, $A, B \in \ell_2$.

y

$$\begin{aligned}
\|A\|^2 &= \sum_{n=0}^{\infty}\left(\left(\frac{\sqrt{2}}{2}\right)^2\right)^n \\
&= \sum_{n=0}^{\infty}\left(\frac{2}{4}\right)^n \\
&= \sum_{n=0}^{\infty}\left(\frac{1}{2}\right)^n \\
&= \frac{1}{1-\frac{1}{2}} \\
&= 2,
\end{aligned}$$

$$\begin{aligned}
\|B\|^2 &= \sum_{n=0}^{\infty}\left(\left(\frac{\sqrt{2}}{10}\right)^2\right)^n \\
&= \sum_{n=0}^{\infty}\left(\frac{2}{10^2}\right)^n \\
&= \sum_{n=0}^{\infty}\left(\frac{1}{50}\right)^n \\
&= \frac{1}{1-\frac{1}{50}} \\
&= \frac{50}{49}\,.
\end{aligned}$$

Por tanto,

$$\begin{aligned}
\theta &= \arccos\left(\frac{\frac{10}{9}}{\sqrt{2}\sqrt{\frac{50}{49}}}\right) \\
&= \arccos\left(\frac{10/9}{10/7}\right) \\
&= \arccos\left(\frac{7}{9}\right) \\
&\approx .679673\,8189 \text{ rad} \\
&\approx \left(\frac{.679673\,8189\cdot 180}{\pi}\right)^{\circ} \\
&\approx 39^{\circ}.
\end{aligned}$$ ◀

❍ **Nota 4.5** Es evidente, de la definición 4.7, que el ángulo entre dos vectores $\vec{u}$ y $\vec{v}$ es $\theta = 90^{\circ}$, si y sólo si $\left(\frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{u}\|\,\|\vec{v}\|}\right) = 0$ si y sólo si $\langle\vec{u},\vec{v}\rangle = 0$; lo cual justifica el concepto de ortogonalidad dado en la definición 4.3 de la subsección 4.1.2 (cfr. pág. 247).

4.1.4 Proyecciones, proceso de ortogonalización, factorización QR

Bases ortonormales

Definición 4.8 *Sea $\mathscr{B}=\{\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_n\}$ una base de un espacio vectorial con producto interior. Se dice que $\mathscr{B}$ es una* ***base ortonormal*** *si*

1. *Los vectores $\vec{u}_i$ son ortogonales entre sí.*
2. $\|\vec{u}_i\|=1\ \forall i=1,2,\ldots n$.

Esto es,

$$\langle u_i,u_j\rangle=\begin{cases}1 & si\ i=j,\\ 0 & si\ i\neq j.\end{cases}$$

► **Ejemplo 4.25** En $\mathbb{R}^3$, $\mathscr{B}=\{(1,0,0),(0,1,0),(0,0,1)\}$ es claramente una base ortonormal.◄

► **Ejemplo 4.26** En $\mathfrak{M}_{2\times2}$, $\mathscr{B}=\{A_{kl}\}$, $k,l=1,2$, donde $A_{kl}=[a_{ij}]$ y

$$a_{ij}=\begin{cases}1 & \text{si } i=k, j=l;\\ 0 & \text{en otro caso,}\end{cases}$$

es una base ortonormal de $\mathfrak{M}_{2\times2}$ como fácilmente puede verificar el lector.◄

Cuando se tiene un conjunto de vectores no nulos ortogonales entre sí, estos vectores son L.I., como se prueba en el siguiente teorema.

Teorema 4.7 *Sean $\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_m$, vectores no nulos ortogonales entre sí. Entonces estos vectores son L.I.*

DEMOSTRACIÓN ■ Sean α_i, $i=1,2,\ldots,m$, números reales tales que

$$\alpha_1\vec{u}_1+\alpha_2\vec{u}_2+\cdots+\alpha_m\vec{u}_m=\vec{0}_{\mathbf{E}}.$$

Entonces, para cada $k=1,2,\ldots,m$,

$$\langle\alpha_1\vec{u}_1+\alpha_2\vec{u}_2+\cdots+\alpha_n\vec{u}_m,\vec{u}_k\rangle=\langle\vec{0}_{\mathbf{E}},\vec{u}_k\rangle.$$

Por tanto,

$$\alpha_1\langle\vec{u}_1,\vec{u}_k\rangle+\alpha_2\langle\vec{u}_2,\vec{u}_k\rangle+\cdots+\alpha_k\langle\vec{u}_k,\vec{u}_k\rangle+\cdots+\alpha_n\langle\vec{u}_m,\vec{u}_k\rangle=0$$

Ya que los vectores son ortogonales entre sí, se sigue que $\langle\vec{u}_i,\vec{u}_k\rangle=0$ si $i\neq k$; y entonces

$$\alpha_k\langle\vec{u}_k,\vec{u}_k\rangle=0.$$

Puesto que ningún vector es nulo, se debe tener $\alpha_k = 0$ para cada $k = 1, 2, \ldots, m$. Así que los vectores $\vec{u}_i$ son L.I. ■

Supongamos ahora que $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ es una base ortogonal de $\mathbf{E}$; es decir, los vectores $\vec{u}_i$ son ortogonales entre sí, entonces si $\vec{v} \in \mathbf{E}$, existen constantes a_i tales que $\vec{v} = \sum_{i=1}^{n} a_i \vec{u}_i$. Luego, para cada $k = 1, 2, \ldots, n$, se tiene

$$\begin{aligned} \langle \vec{v}, \vec{u}_k \rangle &= \left\langle \sum_{i=1}^{n} a_i \vec{u}_i, \vec{u}_k \right\rangle \\ &= \sum_{i=1}^{n} a_i \langle \vec{u}_i, \vec{u}_k \rangle \\ &= a_k \langle \vec{u}_k, \vec{u}_k \rangle \\ &= a_k \|\vec{u}_k\|^2 . \end{aligned}$$

Así que,

$$a_k = \frac{\langle \vec{v}, \vec{u}_k \rangle}{\|\vec{u}_k\|^2}$$

para cada $k = 1, 2, \ldots, n$. Así hemos probado el siguiente teorema.

Teorema 4.8 *Sea $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ una base del espacio $\mathbf{E}$. Entonces,*

1. *Si los vectores son ortogonales entre sí y $\vec{v} \in \mathbf{E}$, se tiene*

$$\vec{v} = \sum_{k=1}^{n} \frac{\langle \vec{v}, \vec{u}_k \rangle}{\|\vec{u}_k\|^2} \vec{u}_k \tag{4.10}$$

2. *Si esta base es ortonormal,*

$$\vec{v} = \sum_{k=1}^{n} \langle \vec{v}, \vec{u}_k \rangle \vec{u}_k \tag{4.11}$$

para todo $\vec{v} \in \mathbf{E}$.

► **Ejemplo 4.27** Es fácil mostrar que $A_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $A_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $A_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $A_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $A_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ y $A_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; esto es, $A_{kl} = [a_{ij}]$, donde

$$a_{ij} = \begin{cases} 1 & \text{si } i = k \text{ y } j = l, \\ 0 & \text{en otro caso;} \end{cases}$$

es una base ortonormal para $\mathfrak{M}_{2\times 3}$. Si $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \in \mathfrak{M}_{2\times 3}$, entonces,

$$\langle B, A_{kl} \rangle = b_{kl}$$

para cada $k = 1,2$ y para cada $l = 1,2,3$; y claramente

$$\begin{aligned} B &= b_{11}A_{11} + b_{12}A_{12} + \cdots + b_{23}A_{23} \\ &= \langle B, A_{11}\rangle A_{11} + \langle B, A_{12}\rangle A_{12} + \cdots + \langle B, A_{23}\rangle A_{23}. \blacktriangleleft \end{aligned}$$

Sea $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ una base ortonormal de $\mathbf{E}$. Por (4.11) del teorema 4.8, todo vector $\vec{v} \in \mathbf{E}$ se puede escribir como

$$\vec{v} = a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_n\vec{u}_n \tag{4.12}$$

donde $a_k = \langle \vec{v}, \vec{u}_k\rangle$, para $k = 1,2,\ldots,n$. Si convenimos en que siempre escribiremos los vectores de $\mathbf{E}$ en esta forma, con los términos en el mismo orden como en (4.12), i.e., el primer término es un escalar por $\vec{u}_1$; el segundo, un escalar por $\vec{u}_2$, etc., y el último término un escalar por $\vec{u}_n$; diremos que esta base ortonormal es **ordenada** y escribiremos $(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$ para enfatizar este hecho; y a $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ lo llamaremos el **vector de coordenadas** del vector $\vec{v}$ relativo a la base ordenada $(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$. Con este convenio, basta conocer el respectivo vector de coordenadas para poder escribir cualquier vector como combinación lineal de los elementos de la base ordenada como en (4.12). Debido a que la base es ortonormal, los cálculos básicos, que involucran al producto interior, se simplifican notablemente y se determinan en función del producto punto en $\mathbb{R}^n$. Hacemos patente este importante resultado en el siguiente teorema, cuya demostración se deja de ejercicio al lector.

Teorema 4.9 *Sea $(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$ una base ordenada ortonormal del espacio $\mathbf{E}$ y sean $(a_1, a_2, \ldots, a_n)$, $(b_1, b_2, \ldots, b_n)$ los vectores de coordenadas de sendos vectores $\vec{u}$ y $\vec{v}$ de $\mathbf{E}$. Entonces,*

1. $$\|\vec{u}\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

2. $$\begin{aligned} \mathrm{d}(\vec{u}, \vec{v}) &= \mathrm{d}((a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n)) \\ &= \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}. \end{aligned}$$

3. $$\begin{aligned} \langle \vec{u}, \vec{v}\rangle &= (a_1, a_2, \ldots, a_n) \cdot (b_1, b_2, \ldots, b_n) \\ &= a_1b_1 + a_2b_2 + \cdots + a_nb_n. \end{aligned}$$

Proyección de un vector sobre un subespacio

En la definición 4.5, establecimos lo que es el vector proyección de un elemento del espacio sobre otro elemento. Ahora estudiaremos el concepto de proyección de un vector sobre un subespacio. Como se ilustra en la figura 4-5, el vector proyección $\vec{p}$ de un vector $\vec{u}$ sobre un subespacio S generado por dos vectores $\vec{v}_1$ y $\vec{v}_2$ debe satisfacer:

1. $\vec{p} \in \mathrm{gn}(\vec{v}_1, \vec{v}_2) = S$.
2. $\vec{u} - \vec{p} \perp \vec{v}_1, \vec{v}_2$ (y, por tanto, a todo $v \in S$).

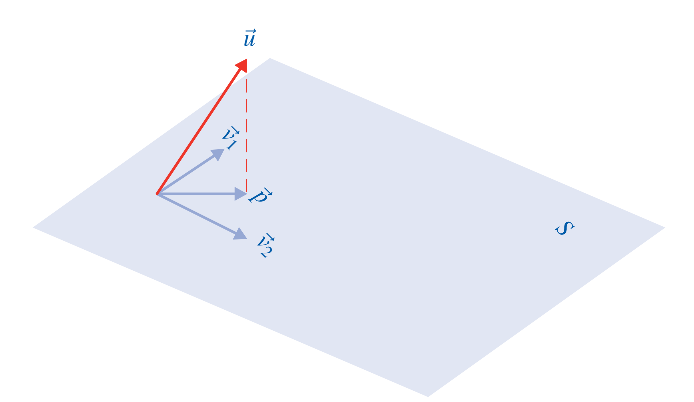


Figura 4-5 • Proyección de un vector $\vec{p}$ sobre un subespacio S en $\mathbb{R}^3$.

En general, tenemos la siguiente definición.

Definición 4.9 *Sea S un subespacio generado por los vectores linealmente independientes $\vec{v}_1$, $\vec{v}_2, \ldots, \vec{v}_m$ y sea $\vec{u} \in \mathbf{E}$. Entonces, al vector $\vec{p}$ que satisface las dos condiciones siguientes:*

1. $\vec{p} \in S$,
2. $\vec{u} - \vec{p} \perp \vec{v}_i$ *para cada* $i = 1, 2, \ldots, m$ *(y, por tanto,* $\vec{u} - \vec{p} \perp \vec{v}\ \forall \vec{v} \in S$*),*

*se le llama el vector **proyección** de $\vec{u}$ sobre el subespacio S.*

Entonces, si $\vec{p}$ es el vector proyección de $\vec{u}$ sobre S, dado que $\vec{p} \in S$, existen $a_i \in \mathbb{R}$ tales que

$$\vec{p} = \sum_{j=1}^{m} a_j \vec{v}_j$$

y, por la condición 2 de la precedente definición, se debe tener

$$\left\langle \vec{u} - \sum_{j=1}^{m} a_j \vec{v}_j, \vec{v}_k \right\rangle = 0$$

para cada $k = 1, 2, \ldots, m$. Esto es,

$$\langle \vec{u}, \vec{v}_k \rangle - \sum_{j=1}^{m} a_j \langle \vec{v}_j, \vec{v}_k \rangle = 0$$

para todo $k = 1, 2, \ldots, m$. Así, para determinar las constantes a_i y, por tanto, el vector $\vec{p}$, se debe resolver el sistema

$$\begin{array}{ccccccccc}
\langle \vec{v}_1, \vec{v}_1 \rangle a_1 & + & \langle \vec{v}_2, \vec{v}_1 \rangle a_2 & + & \cdots & + & \langle \vec{v}_m, \vec{v}_1 \rangle a_m & = & \langle \vec{u}, \vec{v}_1 \rangle \\
\langle \vec{v}_1, \vec{v}_2 \rangle a_1 & + & \langle \vec{v}_2, \vec{v}_2 \rangle a_2 & + & \cdots & + & \langle \vec{v}_m, \vec{v}_2 \rangle a_m & = & \langle \vec{u}, \vec{v}_2 \rangle \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\langle \vec{v}_1, \vec{v}_m \rangle a_1 & + & \langle \vec{v}_2, \vec{v}_m \rangle a_2 & + & \cdots & + & \langle \vec{v}_m, \vec{v}_m \rangle a_m & = & \langle \vec{u}, \vec{v}_m \rangle
\end{array} \tag{4.13}$$

▶ **Ejemplo 4.28** En $C[0,1]$, sean $f(x)=1$ y $g(x)=x$; $S=\text{gn}(f,g)$ y $u(x)=x^2$. Entonces, por la definición precedente, el vector proyección de u sobre S debe tener la forma $p(x)=bx+a\in S$ y $\langle u-p,f\rangle=0=\langle u-p,g\rangle$; esto es,

$$\begin{aligned} 0 &= \int_0^1 \left(x^2-bx-a\right)dx \\ &= \frac{1}{3}-a-\frac{1}{2}b; \end{aligned}$$

$$\begin{aligned} 0 &= \int_0^1 \left(x^3-bx^2-ax\right)dx \\ &= \frac{1}{4}-\frac{1}{3}b-\frac{1}{2}a\,. \end{aligned}$$

Se tiene entonces el sistema

$$\begin{aligned} \frac{1}{3}-a-\frac{1}{2}b &= 0 \\ \frac{1}{4}-\frac{1}{3}b-\frac{1}{2}a &= 0 \end{aligned}$$

que al resolver, resulta $a=-\frac{1}{6}$ y $b=1$. Por lo que

$$p(x)=x-\frac{1}{6}$$

es el vector proyección buscado.◀

Aunque en el ejemplo precedente se pudo encontrar la proyección de un vector sobre un subespacio, no queda claro aún si esto es posible en cualquier situación. Afortunadamente el siguiente teorema garantiza que el sistema (4.13) tiene solución única en todos los casos cuando los vectores $\vec{v}_i$ son linealmente independientes (L.I.).

Teorema 4.10 *Sean $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_m$ vectores L.I. en* **E***, entonces las filas de la matriz simétrica $A=[\langle\vec{v}_i,\vec{v}_j\rangle]$ (y por tanto las columnas) son vectores L.I. de $\mathbb{R}^m$. Por ende, la matriz $[\langle\vec{v}_i,\vec{v}_j\rangle]\in\mathfrak{M}_{m\times m}$ es invertible.*

DEMOSTRACIÓN ■ Sean $\vec{F}_i=(\langle\vec{v}_i,\vec{v}_1\rangle,\langle\vec{v}_i,\vec{v}_2\rangle,\ldots,\langle\vec{v}_i,\vec{v}_m\rangle)$ las m filas de la matriz A y $\alpha_1,\alpha_2,\ldots,\alpha_m\in\mathbb{R}$ tales que

$$\alpha_1\vec{F}_1+\alpha_2\vec{F}_2+\cdots+\alpha_m\vec{F}_m=\vec{0}_{\mathbb{R}^m}.$$

Si $i=1,2,\ldots,m$, se tiene

$$\begin{aligned} \alpha_i\vec{F}_i &= (\alpha_i\langle\vec{v}_i,\vec{v}_1\rangle,\alpha_i\langle\vec{v}_i,\vec{v}_2\rangle,\ldots,\alpha_i\langle\vec{v}_i,\vec{v}_m\rangle) \\ &= (\langle\alpha_i\vec{v}_i,\vec{v}_1\rangle,\langle\alpha_i\vec{v}_i,\vec{v}_2\rangle,\ldots,\langle\alpha_i\vec{v}_i,\vec{v}_m\rangle)\,. \end{aligned}$$

Entonces,

$$\begin{aligned} \vec{0}_{\mathbb{R}^m} &= \alpha_1\vec{F}_1+\alpha_2\vec{F}_2+\cdots+\alpha_m\vec{F}_m \\ &= \left(\sum_{i=1}^m\langle\alpha_i\vec{v}_i,v_1\rangle,\sum_{i=1}^m\langle\alpha_i\vec{v}_i,\vec{v}_2\rangle,\ldots,\sum_{i=1}^m\langle\alpha_i\vec{v}_i,\vec{v}_m\rangle\right). \end{aligned}$$

Por tanto,

$$\begin{array}{rcl} \sum_{i=1}^{m}\langle\alpha_i\vec{v}_i,\vec{v}_1\rangle & = & 0\,, \\ \sum_{i=1}^{m}\langle\alpha_i\vec{v}_i,\vec{v}_2\rangle & = & 0\,, \\ \vdots & \vdots & \vdots \\ \sum_{i=1}^{m}\langle\alpha_i\vec{v}_i,\vec{v}_m\rangle & = & 0\,. \end{array}$$

Que, por propiedades del producto interior, se transforman en

$$\begin{array}{rcl} \langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\vec{v}_1\rangle & = & 0\,, \\ \langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\vec{v}_2\rangle & = & 0\,, \\ \vdots & \vdots & \vdots \\ \langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\vec{v}_m\rangle & = & 0\,. \end{array}$$

Entonces,

$$\begin{array}{rcl} \alpha_1\langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\vec{v}_1\rangle & = & 0\,, \\ \alpha_2\langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\vec{v}_2\rangle & = & 0\,, \\ \vdots & \vdots & \vdots \\ \alpha_m\langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\vec{v}_m\rangle & = & 0\,. \end{array}$$

Nuevamente, por propiedades del producto interior se tiene:

$$\begin{array}{rcl} \langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\alpha_1\vec{v}_1\rangle & = & 0\,, \\ \langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\alpha_2\vec{v}_2\rangle & = & 0\,, \\ \vdots & \vdots & \vdots \\ \langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\alpha_m\vec{v}_m\rangle & = & 0\,. \end{array}$$

Luego,

$$\langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\alpha_1\vec{v}_1\rangle+\langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\alpha_2\vec{v}_2\rangle+\cdots+\langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\alpha_m\vec{v}_m\rangle=0.$$

Que, por propiedades del producto interior, equivale a

$$\langle\sum_{i=1}^{m}\alpha_i\vec{v}_i,\sum_{i=1}^{m}\alpha_i\vec{v}_i\rangle=0\,.$$

Por tanto,

$$\sum_{i=1}^{m}\alpha_i\vec{v}_i=\vec{0}_{\mathbf{E}}\,.$$

Por ser los vectores $\vec{v}_i$ L.I., se deduce que $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$; en consecuencia, las filas $\vec{F}_i$ son vectores L.I. ■

Tenemos como corolario a este teorema que el sistema (4.13) tiene solución única cuando los vectores son L.I.; esto es, si los vectores $\vec{v}_i$ son L.I. y $\vec{u} \in \mathbf{E}$, entonces existe y además es única la proyección de $\vec{u}$ sobre el subespacio S generado por estos vectores. Además, si los vectores $\vec{v}_i$ forman un conjunto ortonormal, los elementos de la matriz $[\langle \vec{v}_i, \vec{v}_j \rangle]$ son nulos cuando $i \neq j$; es decir, la única solución a este sistema está dada por $a_i = \langle \vec{u}, \vec{v}_i \rangle$; luego

$$\vec{p} = \sum_{j=1}^{m} \langle \vec{u}, \vec{v}_j \rangle \, \vec{v}_j$$

es la proyección de $\vec{u}$ sobre S. Si los elementos de la base son sólo ortogonales (no necesariamente unitarios; i.e., de norma uno), entonces la única solución del sistema (4.13) está dada por $a_i = \langle \vec{u}, \vec{v}_i \rangle \,/\, \|\vec{v}_i\|^2$; así que en este caso

$$\vec{p} = \sum_{j=1}^{m} \frac{\langle \vec{u}, \vec{v}_j \rangle}{\|\vec{v}_j\|^2} \vec{v}_j \, .$$

Resumimos lo precedente en el siguiente teorema.

Teorema 4.11 *Sean S un subespacio vectorial de $\mathbf{E}$, $\vec{u} \in \mathbf{E}$ y $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ una base de S. Entonces existe y es único el vector proyección de $\vec{u}$ sobre S. Además:*

1. *$\vec{p}$ está dado por*

$$\vec{p} = \sum_{i=1}^{m} a_i \vec{v}_i$$

donde $(a_1, a_2, \ldots, a_m)$ es la única solución del sistema (4.13).

2. *Si la base $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ es ortonormal,*

$$\vec{p} = \sum_{j=1}^{m} \langle \vec{u}, \vec{v}_j \rangle \, \vec{v}_j \, . \tag{4.14}$$

3. *Si la base $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ sólo es ortogonal,*

$$\vec{p} = \sum_{j=1}^{m} \frac{\langle \vec{u}, \vec{v}_j \rangle}{\|\vec{v}_j\|^2} \vec{v}_j \, . \tag{4.15}$$

► **Ejemplo 4.29** En el ejemplo 4.11 se mostró que las funciones $\operatorname{sen}(x)$ y $\cos(x)$ son ortogonales en el espacio $C[0,\pi]$. Si $g(x) = x$, hallar la proyección p de g sobre $S = \operatorname{gn}(\operatorname{sen}(x), \cos(x))$. ◄

Solución Por (4.15),

$$p(x) = \frac{\langle x, \operatorname{sen}(x) \rangle}{\|\operatorname{sen}(x)\|^2} \operatorname{sen}(x) + \frac{\langle x, \cos(x) \rangle}{\|\cos(x)\|^2} \cos(x).$$

En el ejemplo 4.14 vimos que

$$\|\text{sen}(x)\| = \left(\int_0^{\pi} \text{sen}^2(x)dx\right)^{1/2} = \sqrt{\frac{\pi}{2}}\,.$$

Por otra parte,

$$\begin{aligned}\|\cos(x)\| &= \left(\int_0^{\pi} \cos^2(x)dx\right)^{1/2}\\ &= \left(\int_0^{\pi} \frac{1+\cos(2x)}{2}dx\right)^{1/2}\\ &= \left(\frac{x}{2}\Big|_0^{\pi} + \frac{1}{4}\,\text{sen}(2x)\Big|_0^{\pi}\right)^{1/2}\\ &= \sqrt{\frac{\pi}{2}}\,;\end{aligned}$$

$$\begin{aligned}\langle x, \text{sen}(x)\rangle &= \int_0^{\pi} x\,\text{sen}(x)dx\\ &= -x\cos(x)|_0^{\pi} + \int_0^{\pi}\cos(x)dx\\ &= \pi + \text{sen}(x)|_0^{\pi}\\ &= \pi\,;\end{aligned}$$

$$\begin{aligned}\int_0^{\pi} x\cos(x)dx &= x\,\text{sen}(x)|_0^{\pi} - \int_0^{\pi}\text{sen}(x)dx\\ &= 0 + \cos(x)|_0^{\pi}\\ &= -2.\end{aligned}$$

Luego,

$$\begin{aligned}p(x) &= \frac{\pi}{\frac{\pi}{2}}\,\text{sen}(x) + \frac{-2}{\frac{\pi}{2}}\cos(x)\\ &= 2\,\text{sen}(x) - \frac{4}{\pi}\cos(x)\end{aligned}$$

es la función proyección de $g(x) = x$ sobre el espacio $S = \text{gn}(\text{sen}(x), \cos(x))$. ✔

Proyecciones en subespacios de $\mathbb{R}^n$ y matriz de proyección

Proyectar un vector sobre un subespacio es particularmente sencillo cuando se hace en el espacio $\mathbb{R}^n$; además, el proceso involucra un formato matricial muy simple que será de gran utilidad para este fin y en una de las aplicaciones más importantes que existen de este tema: el método de mínimos cuadrados, el cual veremos más adelante.

Sean $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k \in \mathbb{R}^n$ vectores L.I. Si $W = \text{gn}(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k)$ y $A \in \mathfrak{M}_{n\times k}$ es la matriz cuyas columnas son $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$, entonces $W = \{A\vec{x}\,|\,\vec{x} \in \mathbb{R}^k\}$. Efectivamente, si $\vec{u} \in \text{gn}(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k)$, para algunos $x_1, x_2, \ldots, x_k \in \mathbb{R}$ se tiene

$$\begin{aligned}\vec{u} &= x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_k\vec{a}_k \\ &= \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_k \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \\ &= A\vec{x}\end{aligned}$$

con[13] $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}$. En particular, el vector proyección $\vec{p}$ de $\vec{b}$ sobre W tiene la forma $\vec{p} = A\vec{r}$, con $\vec{r} = (r_1, r_2, \ldots, r_k)$. Además, $\vec{b} - \vec{p} \perp \vec{u}$, $\forall \vec{u} \in W$. Por tanto, $\vec{u} \cdot (\vec{b} - \vec{p}) = 0\ \forall \vec{u} \in W$. Entonces, $\forall \vec{x} \in \mathbb{R}^k$:

$$\begin{aligned}(A\vec{x})^t(\vec{b} - A\vec{r}) = 0 \Rightarrow \\ \vec{x}^t A^t(\vec{b} - A\vec{r}) = 0 \Rightarrow \\ \vec{x}^t(A^t\vec{b} - A^tA\vec{r}) = 0.\end{aligned}$$

Es decir, el producto punto del vector $A^t\vec{b} - A^tA\vec{r}$ con cualquier vector $\vec{x}$ de $\mathbb{R}^k$ es cero, lo que implica, por 7 del teorema 4.1, $A^t\vec{b} - A^tA\vec{r} = \vec{0}_{\mathbb{R}^k}$ y, por tanto,

$$A^t b = (A^tA)\vec{r}.$$

Ahora bien, dado que los vectores $\vec{a}_i$ son L.I., la matriz A tiene rango k; sabemos, por el teorema 3.23 (cfr. pág. 171), que los rangos de A y A^tA son iguales, así que $\operatorname{Rang}(A^tA) = k$, y puesto que A^tA es una matriz cuadrada de orden k, se sigue que es invertible. Luego,

$$\vec{r} = (A^tA)^{-1}A^t\vec{b}$$

y $\vec{p} = A\vec{r} = [A(A^tA)^{-1}A^t]\vec{b}$.

En resumen, el **vector proyección**, $\vec{p}$, de $\vec{b}$ sobre $W = \operatorname{gn}(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k)$ está dado por:

$$\vec{p} = [A(A^tA)^{-1}A^t]\vec{b} \tag{4.16}$$

donde A es la matriz que tiene como columnas a los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$. A la matriz

$$A(A^tA)^{-1}A^t \tag{4.17}$$

se le llama la **matriz de proyección** para el subespacio W.

► **Ejemplo 4.30** Hallar la proyección del vector $\vec{b} = (b_1, b_2, b_3)$ sobre el plano $2x - y - 3z = 0$.◄

Solución Claramente la solución de este sistema escalonado es

$$x = \frac{y + 3z}{2}.$$

[13] Cfr. (1.15) en la página 12.

Es decir,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{s+3r}{2} \\ s \\ r \end{bmatrix} = \begin{bmatrix} q+3t \\ 2q \\ 2t \end{bmatrix} = q\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}.$$

Así, si W es el plano $2x - y - 3z = 0$, entonces $W = \text{gn}((1,2,0),(3,0,2))$. Luego,

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{y} \quad A^t = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 2 \end{bmatrix}.$$

Por lo que,

$$\begin{aligned} A^tA &= \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 3 \\ 3 & 13 \end{bmatrix}. \end{aligned}$$

Utilicemos ahora la adjunta para calcular la inversa de la matriz invertible A^tA, $(B^{-1} = (1/|B|)\text{Adj}(B))$:

$$(A^tA)^{-1} = \frac{1}{56}\begin{bmatrix} 13 & -3 \\ -3 & 5 \end{bmatrix}.$$

La matriz de proyección (4.17) es, en este caso,

$$\begin{aligned} A(A^tA)^{-1}A^t &= \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 2 \end{bmatrix}\left(\frac{1}{56}\begin{bmatrix} 13 & -3 \\ -3 & 5 \end{bmatrix}\right)\begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 2 \end{bmatrix} \\ &= \frac{1}{56}\begin{bmatrix} 4 & 12 \\ 26 & -6 \\ -6 & 10 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ 3 & 2 & 0 \end{bmatrix} \\ &= \frac{1}{56}\begin{bmatrix} 40 & 8 & 24 \\ 8 & 52 & -12 \\ 24 & -12 & 20 \end{bmatrix} \\ &= \begin{bmatrix} 5/7 & 1/7 & 3/7 \\ 1/7 & 13/14 & -3/14 \\ 3/7 & -3/14 & 5/14 \end{bmatrix}. \end{aligned}$$

Por lo que el vector proyección (4.16) es

$$\vec{p} = \begin{bmatrix} 5/7 & 1/7 & 3/7 \\ 1/7 & 13/14 & -3/14 \\ 3/7 & -3/14 & 5/14 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \quad ✔$$

Proceso de ortogonalización de Gram-Schmidt

Las relaciones (4.14) y (4.15) del teorema 4.11 proveen, entre otras cosas, una forma muy simple de proyectar un vector sobre un subespacio cuando éste posee una base ortogonal u ortonormal. En el

teorema 4.9 también vimos que las operaciones que involucran producto interior se simplifican con este tipo de bases; por eso, sería muy bueno el poder contar con bases ortogonales en los espacios donde se esté trabajando. Afortunadamente es posible construir bases ortonormales en espacios de dimensión finita mediante el proceso que a continuación explicamos.

Supongamos que $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ es una base de un subespacio S del espacio $\mathbf{E}$. Sea $\vec{u}_1 = \dfrac{1}{\|\vec{v}_1\|}\vec{v}_1$, entonces $\|\vec{u}_1\| = 1$ y $\operatorname{gn}(\vec{u}_1) = \operatorname{gn}(\vec{v}_1) \subset S$. Sea $\vec{p}_1$ el vector proyección de $\vec{v}_2$ sobre $\vec{u}_1$ y $\vec{w}_2 = \vec{v}_2 - \vec{p}_1$, como se indica en la figura 4-6.

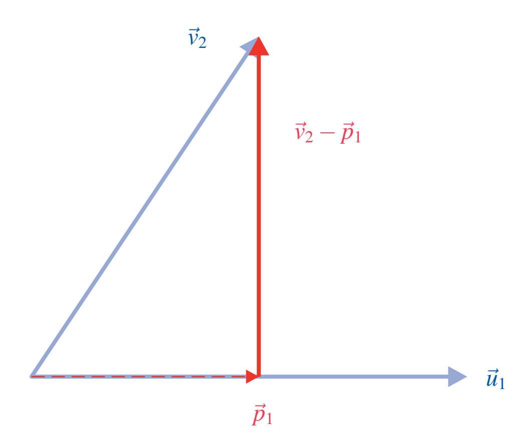


Figura 4-6 • Si $\vec{p}_1$ es el vector proyección de $\vec{v}_2$ sobre $\vec{u}_1$, entonces $\vec{w}_2 = \vec{v}_2 - \vec{p}_1 \perp \vec{u}_1$.

Entonces, $\vec{w}_2 \neq \vec{0}_{\mathbf{E}}$, pues en caso contrario $\vec{v}_2 = \vec{p}_1 \in \operatorname{gn}(\vec{u}_1) = \operatorname{gn}(\vec{v}_1)$ y $\vec{v}_1$, $\vec{v}_2$ serían L.D. Por (4.6) del teorema 4.4 (cfr. pág. 253)

$$\vec{p}_1 = \frac{\langle \vec{v}_2, \vec{u}_1 \rangle}{\|\vec{u}_1\|^2}\vec{u}_1 = \langle \vec{v}_2, \vec{u}_1 \rangle \vec{u}_1$$

y, por ende,

$$\vec{w}_2 = \vec{v}_2 - \langle \vec{v}_2, \vec{u}_1 \rangle \vec{u}_1.$$

Así, $\vec{w}_2 \perp \vec{u}_1$ y $\vec{w}_2 \in \operatorname{gn}(\vec{u}_1, \vec{v}_2) \subset S$; si definimos

$$\vec{u}_2 = \frac{1}{\|\vec{w}_2\|}\vec{w}_2,$$

$\vec{u}_2 \perp \vec{u}_1$, $\|\vec{u}_2\| = 1$ y $\operatorname{gn}(\vec{u}_1, \vec{u}_2) = \operatorname{gn}(\vec{v}_1, \vec{v}_2) \subset S$. Sean ahora $\vec{p}_2$ el vector proyección de $\vec{v}_3$ sobre el subespacio $S_2 = \operatorname{gn}(\vec{u}_1, \vec{u}_2)$ y $\vec{w}_3 = \vec{v}_3 - \vec{p}_2$, como se ilustra en la figura 4-7. Entonces $\vec{w}_3 \neq \vec{0}_{\mathbf{E}}$, ya que en caso contrario

$$\vec{v}_3 = \vec{p}_2 \in \operatorname{gn}(\vec{u}_1, \vec{u}_2) = \operatorname{gn}(\vec{v}_1, \vec{v}_2)$$

y $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ serían L.D. Puesto que $\{\vec{u}_1, \vec{u}_2\}$ es una base ortonormal de S, por (4.14) del teorema 4.11 (cfr. pág. 269), $\vec{p}_2 = \langle \vec{v}_3, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{v}_3, \vec{u}_2 \rangle \vec{u}_2$; luego,

$$\vec{w}_3 = \vec{v}_3 - \langle \vec{v}_3, \vec{u}_1 \rangle \vec{u}_1 - \langle \vec{v}_3, \vec{u}_2 \rangle \vec{u}_2$$

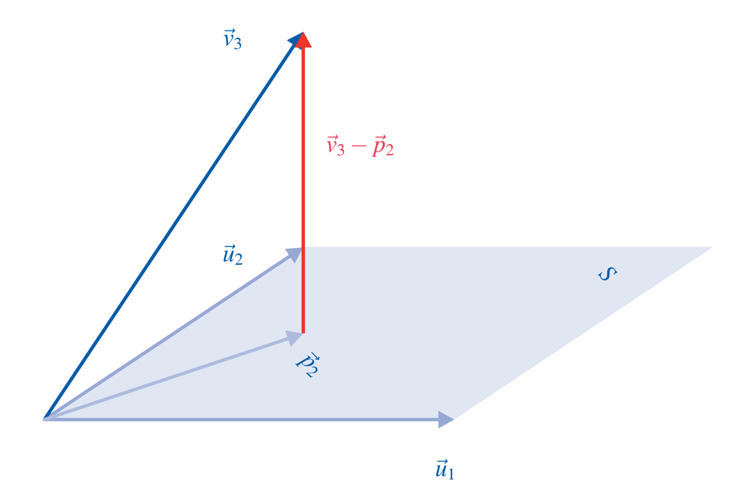


Figura 4-7 • Si $\vec{p}_2$ es la proyección de $\vec{v}_3$ sobre $\text{gn}(\vec{u}_1, \vec{u}_2)$, entonces $\vec{w}_3 = \vec{v}_3 - \vec{p}_2$ es ortogonal a $\vec{u}_1$ y a $\vec{u}_2$.

satisface: $\vec{w}_3 \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{v}_3) \subset S$, $\vec{w}_3 \perp \vec{u}_1$, $\vec{w}_3 \perp \vec{u}_2$. Por tanto, si definimos

$$\vec{u}_3 = \frac{1}{\|\vec{w}_3\|}\vec{w}_3,$$

se tiene: $\text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3) = \text{gn}(\vec{v}_1, \vec{v}_2, \vec{v}_3) \subset S$, $\|\vec{u}_3\| = 1$, y los vectores $\vec{u}_1, \vec{u}_2, \vec{u}_3$ son ortogonales entre sí.

Supongamos que hemos construido k vectores $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$ mediante este proceso, que satisfacen:

1. Los vectores son ortogonales entre sí.
2. Los vectores son unitarios; es decir, $\|\vec{u}_i\| = 1$ para $i = 1, 2, \dots, k$.
3. $\text{gn}(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k) = \text{gn}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) \subset S$.

Sea $\vec{p}_k$ el vector proyección de $\vec{v}_{k+1}$ sobre el subespacio $\text{gn}(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k)$ y $\vec{w}_{k+1} = \vec{v}_{k+1} - \vec{p}_k$; entonces, $\vec{w}_{k+1} \neq \vec{0}_{\mathbf{E}}$, ya que en caso contrario

$$\vec{v}_{k+1} = \vec{p}_k \in \text{gn}(\vec{u}_1, \dots, \vec{u}_k) = \text{gn}(\vec{v}_1, \dots, \vec{v}_k)$$

y $\vec{v}_1, \dots, \vec{v}_k, \vec{v}_{k+1}$ serían L.D. Por (4.14) (dado que $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ es ortonormal),

$$\vec{p}_k = \langle \vec{v}_{k+1}, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{v}_{k+1}, \vec{u}_2 \rangle \vec{u}_2 + \cdots + \langle \vec{v}_{k+1}, \vec{u}_k \rangle \vec{u}_k$$

y, por tanto,

$$\vec{w}_{k+1} = \vec{v}_{k+1} - \langle \vec{v}_{k+1}, \vec{u}_1 \rangle \vec{u}_1 - \langle \vec{v}_{k+1}, \vec{u}_2 \rangle \vec{u}_2 - \cdots - \langle \vec{v}_{k+1}, \vec{u}_k \rangle \vec{u}_k$$

satisface: $\vec{w}_{k+1} \in \text{gn}(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k, \vec{v}_{k+1}) \subset S$ y los vectores $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k, \vec{w}_{k+1}$ son ortogonales entre sí. Por tanto, si definimos

$$\vec{u}_{k+1} = \frac{1}{\|\vec{w}_{k+1}\|}\vec{w}_{k+1}$$

se tiene:

1. Los vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \vec{u}_{k+1}$ son ortogonales entre sí.
2. Los vectores son unitarios ($\|\vec{u}_i\| = 1 \quad \forall i = 1, 2, \ldots, k+1$).
3. $\operatorname{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \vec{u}_{k+1}) = \operatorname{gn}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k, \vec{v}_{k+1}) \subset S$.

Podemos continuar este proceso hasta que $k = m-1$. Así tenemos que los vectores $\vec{u}_i$, $i = 1, 2, \ldots, m$, son ortogonales entre sí, unitarios y $\operatorname{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m) \subset S$. Puesto que el conjunto $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\}$ es ortonormal, estos vectores son L.I.; y ya que $\dim(S) = m$, se tiene que $\operatorname{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m) = S$. Luego $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\}$ es una base ortonormal para S. Este es el llamado proceso de ortogonalización de Gram-Schmidt para construir una base ortonormal de un subespacio de dimensión finita a partir de una base dada de este subespacio. A fin de simplificar la información resumimos este proceso en el siguiente teorema, el cual ya hemos demostrado en las precedentes líneas.

Teorema 4.12 ***(Proceso de ortogonalización de Gram-Schmidt)*** *Sean S un subespacio del espacio* **E** *y $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ una base de S. Entonces, los vectores $\vec{u}_i$, $i = 1, 2, \ldots, m$, definidos en forma recurrente mediante el siguiente proceso:*

$$\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 \tag{4.18}$$

$$\vec{w}_{k+1} = \vec{v}_{k+1} - \langle \vec{v}_{k+1}, \vec{u}_1 \rangle \vec{u}_1 - \langle \vec{v}_{k+1}, \vec{u}_2 \rangle \vec{u}_2 - \cdots - \langle \vec{v}_{k+1}, \vec{u}_k \rangle \vec{u}_k \tag{4.19}$$

$$\vec{u}_{k+1} = \frac{1}{\|\vec{w}_{k+1}\|}\vec{w}_{k+1}, \quad k = 1, 2, \ldots, m-1 \tag{4.20}$$

forman una base ortonormal para S.

► **Ejemplo 4.31** Construir una base ortonormal para $\mathbb{R}^3$ a partir de la base

$$\{(1,1,1), (0,1,1), (0,0,1)\}. \blacktriangleleft$$

Solución

$$\begin{aligned} \vec{u}_1 &= \frac{(1,1,1)}{\sqrt{3}} \quad \text{(por (4.18) del teorema 4.12)} \\ &= (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}). \end{aligned}$$

$$\begin{aligned} \vec{w}_2 &= \vec{v}_2 - \langle \vec{v}_2, \vec{u}_1 \rangle \vec{u}_1 \quad \text{(por (4.19) del teorema 4.12 con } k = 1) \\ &= (0,1,1) - \langle (0,1,1), (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}) \rangle (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}) \\ &= (0,1,1) - (2/\sqrt{3})(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}) \\ &= (0,1,1) - (2/3, 2/3, 2/3) \\ &= (-2/3, 1/3, 1/3). \end{aligned}$$

$$\begin{aligned}\vec{u}_2 &= \frac{\vec{w}_2}{\|\vec{w}_2\|} \quad \text{(por (4.20) del teorema 4.12 con } k=1)\\ &= \frac{(-2/3,1/3,1/3)}{\sqrt{4/9+1/9+1/9}}\\ &= (3/\sqrt{6})(-2/3,1/3,1/3)\\ &= (-2/\sqrt{6},1/\sqrt{6},1/\sqrt{6}).\end{aligned}$$

$$\begin{aligned}\vec{w}_3 &= \vec{v}_3-\langle\vec{v}_3,\vec{u}_1\rangle\vec{u}_1-\langle\vec{v}_3,\vec{u}_2\rangle\vec{u}_2 \quad \text{(por (4.19) del teorema 4.12 con } k=2)\\ &= (0,0,1)-\langle(0,0,1),(1/\sqrt{3},1/\sqrt{3},1/\sqrt{3})\rangle(1/\sqrt{3},1/\sqrt{3},1/\sqrt{3})\\ &\quad -\langle(0,0,1),(-2/\sqrt{6},1/\sqrt{6},1/\sqrt{6},1/\sqrt{6})\rangle(-2/\sqrt{6},1/\sqrt{6},1/\sqrt{6},)\\ &= (0,0,1)-(1/\sqrt{3})(1/\sqrt{3},1/\sqrt{3},1/\sqrt{3})\\ &\quad -(1/\sqrt{6})(-2/\sqrt{6},1/\sqrt{6},1/\sqrt{6})\\ &= (0,-1/2,1/2).\end{aligned}$$

$$\begin{aligned}\vec{u}_3 &= \frac{\vec{w}_3}{\|\vec{w}_3\|} \quad \text{(por (4.20) del teorema 4.12 con } k=2)\\ &= \frac{(0,-1/2,1/2)}{\sqrt{1/2}}\\ &= (0,-\sqrt{2}/2,\sqrt{2}/2).\end{aligned}$$

La base ortonormal buscada es: $\{\vec{u}_1,\vec{u}_2,\vec{u}_3\}$. ✔

▶ **Ejemplo 4.32** Encontrar una base ortonormal para el subespacio $\text{gn}(e^x,x)$ en $C[0,1]$.◀

Solución En este caso, $\vec{v}_1=e^x$ y $\vec{v}_2=x$; entonces,

$$\begin{aligned}\|\vec{v}_1\| &= \langle\vec{v}_1,\vec{v}_1\rangle^{1/2}\\ &= \left[\int_0^1 e^{2x}dx\right]^{1/2}\\ &= \left[\frac{1}{2}e^{2x}\Big|_0^1\right]^{1/2}\\ &= (1/\sqrt{2})(e^2-1)^{1/2}.\end{aligned}$$

Por (4.18) del teorema 4.12,

$$\vec{u}_1=\frac{\sqrt{2}}{(e^2-1)^{1/2}}e^x.$$

De (4.19) del teorema 4.12, con $k=1$,

$$\vec{w}_2 = x-\left\langle x,\frac{\sqrt{2}}{(e^2-1)^{1/2}}e^x\right\rangle\frac{\sqrt{2}}{(e^2-1)^{1/2}}e^x.$$

Calculemos el producto interior de la última igualdad:

$$\left\langle x, \frac{\sqrt{2}}{(e^2-1)^{1/2}} e^x \right\rangle = \frac{\sqrt{2}}{(e^2-1)^{1/2}} \langle x, e^x \rangle$$

y

$$\begin{aligned} \langle x, e^x \rangle &= \int_0^1 x e^x dx \\ &= x e^x |_0^1 - \int_0^1 e^x dx \\ &= e - [e^x]_0^1 \\ &= 1; \end{aligned}$$

por tanto,

$$\left\langle x, \frac{\sqrt{2}}{(e^2-1)^{1/2}} e^x \right\rangle = \frac{\sqrt{2}}{(e^2-1)^{1/2}}.$$

Así,

$$\begin{aligned} \vec{w}_2 &= x - \frac{\sqrt{2}}{(e^2-1)^{1/2}} \frac{\sqrt{2}}{(e^2-1)^{1/2}} e^x \\ &= x - \frac{2}{e^2-1} e^x. \end{aligned}$$

Para normalizar[14] $\vec{w}_2$ tenemos que

$$\begin{aligned} \|\vec{w}_2\|^2 &= \int_0^1 \left(x - \frac{2}{e^2-1} e^x \right)^2 dx \\ &= \int_0^1 \left[x^2 - \frac{4xe^x}{e^2-1} + \frac{4e^{2x}}{(e^2-1)^2} \right] dx \\ &= \left[\frac{x^3}{3} \right]_0^1 - \frac{4}{e^2-1}(1) + \left[\frac{2}{(e^2-1)^2} e^{2x} \right]_0^1 \\ &= \frac{1}{3} - \frac{4}{e^2-1} + \frac{2}{(e^2-1)^2}(e^2-1) \\ &= \frac{1}{3} - \frac{2}{e^2-1} \\ &= \frac{e^2-7}{3(e^2-1)}; \end{aligned}$$

de donde (cfr. (4.20) del teorema 4.12, con $k = 1$)

$$\vec{u}_2 = \left(\frac{3(e^2-1)}{e^2-7} \right)^{1/2} \left(x - \frac{2}{e^2-1} e^x \right)$$

[14]Emplearemos el término *normalizar* cuando multipliquemos por el recíproco de la norma de un vector para obtener un vector con la misma dirección pero con norma 1.

y la base buscada es:

$$\left\{\frac{\sqrt{2}}{(e^2-1)^{1/2}}e^x, \left(\frac{3(e^2-1)}{e^2-7}\right)^{1/2}\left(x-\frac{2}{e^2-1}e^x\right)\right\}. \checkmark$$

Sabemos, de la discusión que hicimos para la derivación del proceso de Gram-Schmidt, que el vector $\vec{u}_{k+1}$ es ortogonal a los vectores $\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_k$ y que $\text{gn}(\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_k)=\text{gn}(\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k)$; por lo que $\vec{u}_{k+1}$ es ortogonal a los vectores $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k$. También de (4.19) y (4.20) del teorema 4.12, se desprende $\langle\vec{v}_k,\vec{u}_k\rangle=\|\vec{w}_k\|\neq 0$, para cada $k=1,\ldots,m$. Con lo que hemos probado el siguiente corolario.

Corolario 4.4 *Sean $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_m$ vectores L.I. del espacio vectorial* **E** *y sean $\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_m$ los vectores ortonormales que se obtienen de los vectores $\vec{v}_i$ por medio del proceso de ortogonalización de Gram-Schmidt. Entonces, $\vec{u}_{k+1}$ es ortogonal a los vectores $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k$, para cada $k=1,2,\ldots,m-1$. Esto es:*

$$\langle\vec{v}_j,\vec{u}_i\rangle = 0 \text{ si } i>j \tag{4.21}$$

Además

$$\langle\vec{v}_i,\vec{u}_i\rangle \neq 0 \text{ para todo } i \tag{4.22}$$

Factorización QR

Supongamos que A es una matriz $n\times m$ cuyas columnas son vectores $\vec{v}_1$, $\vec{v}_2,\ldots,\vec{v}_m$ linealmente independientes de $\mathbb{R}^n$. Sea $\{\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_m\}$ la base ortonormal que se obtiene por medio del proceso de ortogonalización de Gram-Schmidt aplicado a la base $\{\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_m\}$ del espacio columna de A. Por (4.11) del teorema 4.8 (cfr. pág. 264) se tiene, para cada $j=1.2,\ldots,m$,

$$\vec{v}_j=\langle\vec{v}_j,\vec{u}_1\rangle\,\vec{u}_1+\langle\vec{v}_j,\vec{u}_2\rangle\,\vec{u}_2+\cdots+\langle\vec{v}_j,\vec{u}_m\rangle\,\vec{u}_m.$$

Sea Q la matriz que tiene como columnas a los vectores $\vec{u}_j$; por (1.5) del ejemplo 1.15 (cfr. pág. 12),

$$\vec{v}_j=Q\begin{bmatrix}\langle\vec{v}_j,\vec{u}_1\rangle\\ \langle\vec{v}_j,\vec{u}_2\rangle\\ \vdots\\ \langle\vec{v}_j,\vec{u}_m\rangle\end{bmatrix}.$$

Entonces, si $\vec{c}_j=\begin{bmatrix}\langle\vec{v}_j,\vec{u}_1\rangle\\ \langle\vec{v}_j,\vec{u}_2\rangle\\ \vdots\\ \langle\vec{v}_j,\vec{u}_m\rangle\end{bmatrix}$,

$$\begin{aligned}A&=\begin{bmatrix}\vec{v}_1 & \vec{v}_1 & \cdots & \vec{v}_m\end{bmatrix}\\ &=\begin{bmatrix}Q\vec{c}_1 & Q\vec{c}_2 & \cdots & Q\vec{c}_m\end{bmatrix}.\end{aligned}$$

Así, si R es la matriz que tiene como columnas a los vectores $\vec{c}_j$, por (1.2) del ejemplo 1.14,

$$\begin{aligned} A &= \begin{bmatrix} \vec{v}_1 & \vec{v}_1 & \cdots & \vec{v}_m \end{bmatrix} \\ &= \begin{bmatrix} Q\vec{c}_1 & Q\vec{c}_2 & \cdots & Q\vec{c}_m \end{bmatrix} \\ &= QR. \end{aligned}$$

Ahora bien, por (4.21) del corolario 4.4, $\langle \vec{v}_j, \vec{u}_i \rangle = 0$ para todo $i > j$; luego, las componentes de la matriz $R = [\langle \vec{v}_j, \vec{u}_i \rangle]$ son nulas para $j < i$; es decir, todas las componentes por debajo de diagonal. Así, la matriz $R = [\langle \vec{v}_j, \vec{u}_i \rangle]$ es triangular superior e invertible, pues, por (4.22) $\langle \vec{v}_i, \vec{u}_i \rangle \neq 0$ para todo i. La matriz A se puede factorizar entonces como el producto de una matriz cuyas columnas son vectores ortonormales y una matriz triangular superior no singular. Hacemos patente esta conclusión en el siguiente teorema.

Teorema 4.13 *Sean $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ vectores linealmente independientes en $\mathbb{R}^n$; A la matriz $n \times m$ cuyas columnas son estos vectores y $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\}$ la base ortonormal que se obtiene del proceso de ortogonalización de Gram-Schmidt aplicado a la base $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ del espacio columna de A. Entonces A se puede factorizar como el producto de una matriz Q y una matriz triangular superior invertible R. Específicamente,*

$$A = QR,$$

donde el primer factor es la matriz Q cuyas columnas son los vectores $\vec{u}_i$ y el segundo factor es la matriz triangular superior no singular $R = [\langle \vec{v}_j, \vec{u}_i \rangle]$; esto es,

$$R = \begin{bmatrix} \langle \vec{v}_1, \vec{u}_1 \rangle & \langle \vec{v}_2, \vec{u}_1 \rangle & \langle \vec{v}_3, \vec{u}_1 \rangle & \cdots & \langle \vec{v}_m, \vec{u}_1 \rangle \\ 0 & \langle \vec{v}_2, \vec{u}_2 \rangle & \langle \vec{v}_3, \vec{u}_2 \rangle & \cdots & \langle \vec{v}_m, \vec{u}_2 \rangle \\ 0 & 0 & \langle \vec{v}_3, \vec{u}_3 \rangle & \cdots & \langle \vec{v}_m, \vec{u}_3 \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \langle \vec{v}_m, \vec{u}_m \rangle \end{bmatrix}.$$

▶ **Ejemplo 4.33** Sea

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

1. Mostrar que las columnas de A son L.I.
2. Encontrar una base ortonormal para el espacio columna de la matriz A.
3. Factorizar A como el producto de una matriz Q cuyas columnas sean vectores ortonormales y una matriz R triangular superior.◀

Solución 1. Llevemos A a la forma escalonada mediante el método de Gauss:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Puesto que toda columna tiene pivote, el sistema $A\vec{x} = \vec{0}$ sólo tiene la solución trivial y, por tanto, los vectores son L.I.

2. Dado que las columnas forman una base (ya que son vectores L.I.) del espacio columna, podemos aplicar el proceso de ortogonalización de Gram-Schmidt para obtener una base ortonormal:

$$\begin{aligned} \vec{u}_1 &= \frac{1}{\sqrt{2}}(1,0,1,0) \\ &= \left(1/\sqrt{2},0,1/\sqrt{2},0\right). \end{aligned}$$

$$\begin{aligned} \vec{w}_2 &= (1,1,1,0) - \left((1,1,1,0)\cdot\left(1/\sqrt{2},0,1/\sqrt{2},0\right)\right)\left(1/\sqrt{2},0,1/\sqrt{2},0\right) \\ &= (1,1,1,0) - \left(2/\sqrt{2}\right)\left(1/\sqrt{2},0,1/\sqrt{2},0\right) \\ &= (1,1,1,0) - (1,0,1,0) \\ &= (0,1,0,0) \ \therefore \end{aligned}$$

$$\vec{u}_2 = \frac{1}{\|\vec{w}_2\|}\vec{w}_2 = (0,1,0,0).$$

$$\begin{aligned} \vec{w}_3 &= (1,-1,0,1) - \left((1,-1,0,1)\cdot\left(1/\sqrt{2},0,1/\sqrt{2},0\right)\right)\left(1/\sqrt{2},0,1/\sqrt{2},0\right) \\ &\quad - ((1,-1,0,1)\cdot(0,1,0,0))(0,1,0,0) \\ &= (1,-1,0,1) - \left(1/\sqrt{2}\right)\left(1/\sqrt{2},0,1/\sqrt{2},0\right) - (-1)(0,1,0,0) \\ &= (1,-1,0,1) + (-1/2,0,-1/2,0) + (0,1,0,0) \\ &= (1/2,0,-1/2,1) \ \therefore \end{aligned}$$

$$\vec{u}_3 = \frac{1}{\|\vec{w}_3\|}\vec{w}_3 = \frac{2}{\sqrt{6}}\left(\frac{1}{2},0,-\frac{1}{2},1\right) = \left(\frac{1}{\sqrt{6}},0,-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}}\right).$$

3. Por el teorema 4.13,

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{6}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{6}} \end{bmatrix}$$

y si $\vec{v}_i$ son los vectores columna de A,

$$R = \begin{bmatrix} \langle \vec{v}_1, \vec{u}_1 \rangle & \langle \vec{v}_2, \vec{u}_1 \rangle & \langle \vec{v}_3, \vec{u}_1 \rangle \\ 0 & \langle \vec{v}_2, \vec{u}_2 \rangle & \langle \vec{v}_3, \vec{u}_2 \rangle \\ 0 & 0 & \langle \vec{v}_3, \vec{u}_3 \rangle \end{bmatrix}.$$

Como

$$\langle \vec{v}_1, \vec{u}_1 \rangle = (1,0,1,0) \cdot \left(1/\sqrt{2}, 0, 1/\sqrt{2}, 0\right) = 2/\sqrt{2},$$

$$\langle \vec{v}_2, \vec{u}_1 \rangle = (1,1,1,0) \cdot \left(1/\sqrt{2}, 0, 1/\sqrt{2}, 0\right) = 2/\sqrt{2},$$

$$\langle \vec{v}_3, \vec{u}_1 \rangle = (1,-1,0,1) \cdot \left(1/\sqrt{2}, 0, 1/\sqrt{2}, 0\right) = 1/\sqrt{2},$$

$$\langle \vec{v}_2, \vec{u}_2 \rangle = (1,1,1,0) \cdot (0,1,0,0) = 1,$$

$$\langle \vec{v}_3, \vec{u}_2 \rangle = (1,-1,0,1) \cdot (0,1,0,0) = -1$$

y

$$\langle \vec{v}_3, \vec{u}_3 \rangle = (1,-1,0,1) \left(1/\sqrt{6}, 0, -1/\sqrt{6}, 2/\sqrt{6}\right) = 3/\sqrt{6},$$

entonces,

$$R = \begin{bmatrix} \frac{2}{\sqrt{2}} & \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 1 & -1 \\ 0 & 0 & \frac{3}{\sqrt{6}} \end{bmatrix}.$$

El lector puede verificar que efectivamente

$$A = QR$$

realizando el producto. ✔

Matrices ortogonales

Sea A una matriz cuadrada. Supongamos que las columnas de A son vectores $\vec{K}_i$ ortonormales de $\mathbb{R}^n$. Entonces, si c_{ij} es la componente ij del producto A^tA,

$$c_{ij} = \vec{K}_i^t \vec{K}_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{en otro caso} \end{cases} \tag{4.23}$$

Luego, A^tA es la matriz identidad. Recíprocamente, si A es una matriz cuadrada de orden n tal que $A^tA = I_n$, la identidad de orden n, entonces nuevamente se tiene la relación (4.23) y, por ende, las columnas de esta matriz son vectores ortonormales de $\mathbb{R}^n$. Hemos probado así el siguiente teorema.

Teorema 4.14 *Sea A una matriz cuadrada de orden n. Las siguientes condiciones son equivalentes a pares:*

1. *Las columnas de A son vectores ortonormales de $\mathbb{R}^n$.*
2. $A^tA = I_n$.
3. *A es una matriz invertible y $A^{-1} = A^t$.*

Definición 4.10 *Una matriz cuadrada A de orden n que satisface una de las condiciones del teorema 4.14 (y, por tanto, las otras dos) se llama* ***matriz ortogonal***.[15]

► **Ejemplo 4.34** Sean los vectores

$$\vec{u}_1 = \left(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}\right), \quad \vec{u}_2 = \left(-2/\sqrt{6}, 1/\sqrt{6}, 1/\sqrt{6}\right) \quad \text{y} \quad \vec{u}_3 = \left(0, -\sqrt{2}/2, \sqrt{2}/2\right).$$

Entonces, si

$$A = \begin{bmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -\sqrt{2}/2 \\ 1/\sqrt{3} & 1/\sqrt{6} & \sqrt{2}/2 \end{bmatrix},$$

tenemos

$$A^tA = \begin{bmatrix} \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} \\ -\frac{1}{3}\sqrt{6} & \frac{1}{6}\sqrt{6} & \frac{1}{6}\sqrt{6} \\ 0 & -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -\sqrt{2}/2 \\ 1/\sqrt{3} & 1/\sqrt{6} & \sqrt{2}/2 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Así los vectores $\vec{u}_1, \vec{u}_2$ y $\vec{u}_3$ son ortonormales (cfr. ejemplo 4.31) y la matriz A es ortogonal.◄

❍ **Nota 4.6**

1. Observe que si las columnas de una matriz A son ortonormales, entonces sus filas también y viceversa; pues, $A^tA = I \Leftrightarrow AA^t = I$.
2. En el caso de que A sea una matriz cuadrada, el teorema de factorización QR (4.13) afirma que si las columnas de A son vectores L.I., entonces A se puede factorizar como el producto de una matriz ortogonal con una matriz triangular superior.

[15] Sería más adecuado usar el término ortonormal, pero éste se reserva para bases y no es común su uso en la literatura de álgebra lineal.

4.1.5 Aproximación óptima de un vector por elementos de un subespacio

En la figura 4-8 se ilustra el conocido hecho de geometría elemental que establece que la distancia mínima de un punto a una línea recta o a un plano es la longitud de la línea perpendicular de ese punto a la línea recta o al plano. En términos de lo que en este texto hemos estudiado, esto significa que la distancia mínima de un vector a un subespacio es la norma de la diferencia entre este vector y su proyección sobre el subespacio. En el siguiente teorema mostraremos que este hecho geométrico también es válido en espacios con producto interior. Nuevamente esta generalización no es una simple curiosidad matemática, sino que es de gran importancia en la teoría de aproximación; tema que introduciremos más adelante. La demostración de este teorema se basa en un artificio algebraico y el teorema de Pitágoras en espacios con producto interior; sin embargo, este artificio es completamente intuitivo geométricamente y vale la pena gastar un poco de tiempo en su explicación. En *(b)* de la figura 4-8, el segmento de línea que va del vector proyección $\vec{p}$ al punto $\vec{u}$ es ortogonal a todo punto de S y, por tanto, a toda línea de este plano; así que en particular es ortogonal a la línea recta que pasa por los puntos $\vec{p}$ y $\vec{v}$. Por ende, el segmento de línea recta que va del punto $\vec{u}$ al punto $\vec{v}$ es la hipotenusa del triángulo rectángulo que pasa por los puntos $\vec{u}$, $\vec{p}$ y $\vec{v}$; y como en todo triángulo rectángulo la longitud de la hipotenusa es mayor que las longitudes de los catetos, se tiene $\mathrm{d}(\vec{u},\vec{p}) \leq \mathrm{d}(\vec{u},\vec{v})$.

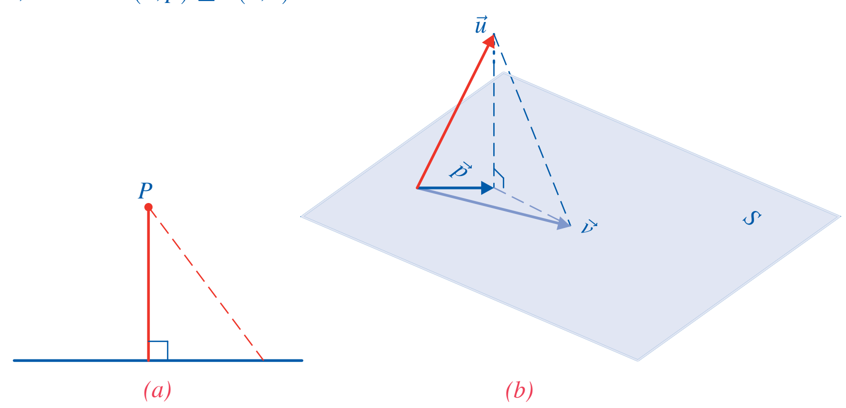


Figura 4-8 • *(a)* La menor distancia de un punto P, exterior a una línea recta, está dada por la longitud de la perpendicular de la línea recta al punto P. *(b)* La menor distancia de un punto $\vec{u}$, exterior a un plano S, a cualquier punto $\vec{v}$ del plano, se alcanza en el vector proyección $\vec{p}$ de $\vec{u}$ sobre S.

Teorema 4.15 *Sean S un subespacio de $\mathbf{E}$ de dimensión finita, $\vec{u} \in \mathbf{E}$ un vector dado y $\vec{p} \in S$ el vector proyección de $\vec{u}$ sobre S. Entonces*

$$\|\vec{u}-\vec{p}\| \leq \|\vec{u}-\vec{v}\| \quad \forall \vec{v} \in S.$$

Es decir, el mínimo de la distancia $\mathrm{d}(\vec{u},\vec{v})$, cuando $\vec{v}$ varía en S, se alcanza en $\vec{v}=\vec{p}$.

DEMOSTRACIÓN ■ Sea $\vec{v} \in S$ cualquier vector, entonces $\vec{p}-\vec{v} \in S$. Por lo que $\vec{u}-\vec{p} \perp \vec{p}-\vec{v}$; así (por el teorema de Pitágoras)

$$\begin{aligned}
\|\vec{u}-\vec{v}\|^2 &= \|(\vec{u}-\vec{p})+(\vec{p}-\vec{v})\|^2 \\
&= \|\vec{u}-\vec{p}\|^2+\|\vec{p}-\vec{v}\|^2 \\
&\geq \|\vec{u}-\vec{p}\|^2;
\end{aligned}$$

de donde

$$\|\vec{u}-\vec{v}\| \geq \|\vec{u}-\vec{p}\|. \quad \blacksquare$$

En forma equivalente, el teorema 4.15 afirma que dados un elemento $\vec{u}$ del espacio $\mathbf{E}$ y S, un subespacio de dimensión finita, existe un elemento $\vec{p} \in S$, de entre todos los elementos de S, que es el que mejor aproxima a $\vec{u}$ en el sentido de la norma inducida por el producto interior; es decir,[16]

$$\min_{\vec{v}\in S} \|\vec{u}-\vec{v}\| = \|\vec{u}-\vec{p}\| \tag{4.24}$$

Además, como el lector seguramente ya intuyó, el vector proyección $\vec{p}$ de $\vec{u}$ sobre S es el único elemento de S que aproxima óptimamente a $\vec{u}$ en S; esto es, cumple con la relación (4.24). En efecto, si $\vec{p}_1 \in S$ y $\|\vec{u}-\vec{p}_1\| \leq \|\vec{u}-\vec{v}\|$ para todo $\vec{v} \in S$, entonces se debe tener

$$\begin{aligned} \|\vec{u}-\vec{p}_1\| &\leq \|\vec{u}-\vec{p}\| \quad \text{y} \\ \|\vec{u}-\vec{p}\| &\leq \|\vec{u}-\vec{p}_1\| \end{aligned}$$

de donde se desprende

$$\|\vec{u}-\vec{p}_1\| = \|\vec{u}-\vec{p}\|.$$

Puesto que $\vec{p}$ es el vector proyección de $\vec{u}$ sobre S, $\vec{u}-\vec{p} \perp \vec{v}$ para todo $\vec{v} \in S$, en particular $\vec{u}-\vec{p} \perp \vec{p}-\vec{p}_1 \in S$. Por el teorema de Pitágoras se tiene

$$\begin{aligned} \|\vec{u}-\vec{p}_1\|^2 &= \|(\vec{u}-\vec{p})+(\vec{p}-\vec{p}_1)\|^2 \\ &= \|\vec{u}-\vec{p}\|^2 + \|\vec{p}-\vec{p}_1\|^2 \\ &= \|\vec{u}-\vec{p}_1\|^2 + \|\vec{p}-\vec{p}_1\|^2; \end{aligned}$$

lo que implica

$$\|\vec{p}-\vec{p}_1\|^2 = 0$$

y, por tanto,

$$\vec{p}_1 = \vec{p}.$$

En resumen, hemos probado que en cualquier espacio con producto interior, $\mathbf{E}$, dados un vector $\vec{u}$ y un subespacio S de dimensión finita, existe un único vector $\vec{p} \in S$ tal que

$$\|\vec{u}-\vec{p}\| \leq \|\vec{u}-\vec{v}\| \quad \forall \vec{v} \in S;$$

es decir, para el cual (4.24) se cumple; donde $\|\cdot\|$ es la norma inducida por el producto interior y, en este caso, el vector $\vec{p}$ es la proyeción de $\vec{u}$ sobre el subespacio S. Llamaremos a $\vec{p}$ la **aproximación óptima** (o la **mejor aproximación**) de $\vec{u}$ en S (relativa a la norma inducida $\|\cdot\|$).

[16] La notación $\min_{\vec{v}\in S} \|\vec{u}-\vec{v}\|$ significa el valor mínimo del conjunto de números reales $\{\|\vec{u}-\vec{v}\| \mid \vec{v} \in S\}$.

❍ **Nota 4.7**

1. En el último apartado de este capítulo estudiaremos el concepto de aproximaciones óptimas, pero en subespacios normados cuya norma no necesariamente proviene de un producto interior.
2. Una norma en el sentido más general, como veremos en la siguiente sección, es una función que permite medir magnitudes de vectores y distancias entre ellos; como se hace con la norma usual en $\mathbb{R}^n$ o con las normas inducidas en espacios con producto interior. Sin embargo, como veremos más adelante, no toda norma proviene de un producto interior.
3. De igual manera a como ocurre en espacios con producto interior veremos, en la siguiente sección, que en un espacio vectorial normado (aunque la norma no provenga de un producto interior), dado un vector $\vec{u}$ y un subespacio de dimensión finita S, siempre existe al menos una aproximación óptima para este vector en S; pero dicha aproximación no necesariamente es única.

A continuación, encontraremos la aproximación óptima de vectores en subespacios para dos casos particulares en el espacio de funciones continuas, con el fin de ilustrar esta importante teoría.

Aproximación con polinomios de Legendre

► **Ejemplo 4.35** Sean $\mathbf{E} = C[-1,1]$ el espacio de funciones continuas en $[-1,1]$ dotado del producto interior

$$\langle f,g\rangle = \int_{-1}^{1} f(x)g(x)dx,$$

$S_2 = \text{gn}\big(1,x,x^2\big)$ y $f \in \mathbf{E}$ una función dada. Apliquemos el proceso de ortogonalización de Gram-Schmidt a la base $\{1,x,x^2\}$:

$$\|1\|^2 = \int_{-1}^{1} dx = x|_{-1}^{1} = 2$$

entonces

$$\begin{aligned}
u_1 &= \frac{1}{\sqrt{2}}.\\
w_2 &= x - \left\langle x, 1/\sqrt{2}\right\rangle 1/\sqrt{2},
\end{aligned}$$

$$\begin{aligned}
\left\langle x, 1/\sqrt{2}\right\rangle 1/\sqrt{2} &= \frac{1}{2}\langle x,1\rangle\\
&= \frac{1}{2}\int_{-1}^{1} xdx\\
&= \frac{1}{2}\,x^2\Big|_{-1}^{1}\\
&= 0,
\end{aligned}$$

$$\begin{aligned}
\|x\|^2 &= \int_{-1}^{1} x^2 dx\\
&= \frac{x^3}{3}\bigg|_{-1}^{1}\\
&= \frac{2}{3};
\end{aligned}$$

así que

$$u_2 = \frac{w_2}{\|w_2\|} = \sqrt{\frac{3}{2}}x.$$

$$w_3 = x^2 - \left\langle x^2, 1/\sqrt{2}\right\rangle 1/\sqrt{2} - \left\langle x^2, \sqrt{3/2}x\right\rangle \sqrt{3/2}x,$$

$$\left\langle x^2, 1/\sqrt{2}\right\rangle 1/\sqrt{2} = \frac{1}{2}\int_{-1}^{1} x^2 dx = \frac{1}{2}\left.\frac{x^3}{3}\right|_{-1}^{1} = \frac{1}{3},$$

$$\left\langle x^2, \sqrt{3/2}x\right\rangle = \sqrt{3/2}\int_{-1}^{1} x^3 dx = \sqrt{3/2}\left.\frac{x^4}{4}\right|_{-1}^{1} = 0;$$

por tanto

$$w_3 = x^2 - \frac{1}{3} \quad \text{y}$$

$$\|w_3\|^2 = \int_{-1}^{1}\left(x^2 - \frac{1}{3}\right)^2 dx = \frac{8}{45};$$

así que

$$u_3 = \frac{1}{\|w_3\|}w_3 = \frac{3}{2}\sqrt{\frac{5}{2}}\left(x^2 - \frac{1}{3}\right).$$

La base ortonormal es

$$\left\{u_1 = \frac{1}{\sqrt{2}}, u_2 = \sqrt{\frac{3}{2}}x, u_3 = \frac{3}{2}\sqrt{\frac{5}{2}}\left(x^2 - \frac{1}{3}\right)\right\}.$$

Entonces, por (4.14) del teorema 4.11 (cfr. pág. 269), la proyección p_2 de f sobre S_2 es

$$p_2 = \langle f, u_1\rangle u_1 + \langle f, u_2\rangle u_2 + \langle f, u_3\rangle u_3.$$

Por ejemplo, si $f(x) = \operatorname{sen}(\pi x)$, entonces[17]

$$\langle f, u_1\rangle = \int_{-1}^{1} \operatorname{sen}(\pi x)\frac{1}{\sqrt{2}}dx = 0,$$
$$\langle f, u_2\rangle = \int_{-1}^{1} \operatorname{sen}(\pi x)\sqrt{\frac{3}{2}}x dx = \frac{1}{\pi}\sqrt{6},$$
$$\langle f, u_2\rangle = \int_{-1}^{1} \operatorname{sen}(\pi x)\frac{3}{2}\sqrt{\frac{5}{2}}\left(x^2 - \frac{1}{3}\right) dx = 0.$$

Luego,

$$p_2(x) = \frac{1}{\pi}\sqrt{6}\sqrt{\frac{3}{2}}x$$
$$= \frac{3}{\pi}x$$

es el polinomio de grado a lo más dos que mejor aproxima a $f(x) = \operatorname{sen}(\pi x)$ en el sentido de que

$$\int_{-1}^{1} (f(x) - p_2(x))^2 \, dx$$

tiene el valor más pequeño al compararla con

$$\int_{-1}^{1} (f(x) - q(x))^2 \, dx$$

para cualquier polinomio $q(x)$ de grado a lo más dos. Las gráficas de $f(x)$ y $p_2(x)$ se ilustran en la figura 4-9. De hecho el lector puede comprobar, realizando la integral, que

$$\left\| \operatorname{sen}(\pi x) - \frac{3}{\pi}x \right\| \approx 0.6261572471.$$

Las gráficas de $(f(x) - p_2(x))^2$ y de $(f(x) - q(x))^2$, para $q(x) = (1/10)\left(x^2 - 2x + 2\right)$, se ilustran en la figura 4-10; en ella se puede apreciar que aparentemente el área bajo la curva $(f(x) - p_2(x))^2$ es menor

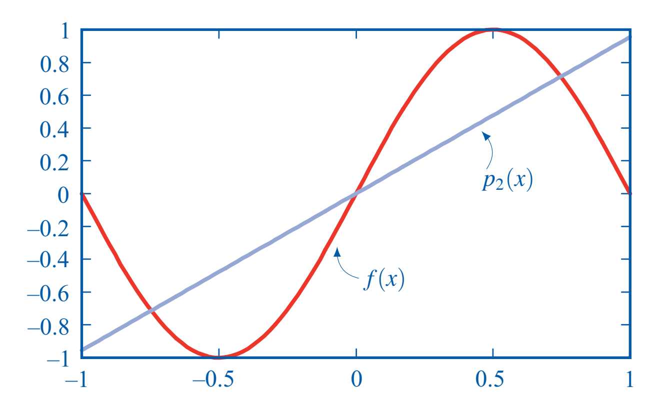


Figura 4-9 • Gráficas de $f(x) = \operatorname{sen}(\pi x)$ y la proyección, $p_2(x)$, de esta función sobre el subespacio $S_2 = \operatorname{gn}(1, x, x^2)$.

[17] Dejamos los detalles de los cálculos de las integrales de ejercicio para el lector.

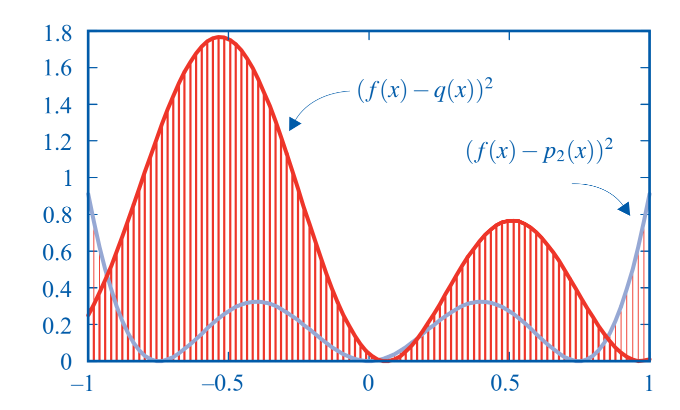


Figura 4-10 • Gráficas de las funciones $(f(x)-p_2(x))^2$ y $(f(x)-q(x))^2$. El área bajo la primera curva debe ser inferior al área bajo la segunda curva.

que el área bajo la curva $(f(x)-q(x))^2$ en $[-1,1]$. Efectivamente,

$$\begin{aligned}\int_{-1}^{1}(\operatorname{sen}(\pi x)-q(x))^2\,dx &= \frac{1}{750\pi}(600+853\pi)\\ &\approx 1.39198\,1242\end{aligned}$$

de donde[18]

$$\|\operatorname{sen}(\pi x)-q(x)\|\approx\sqrt{1.39198\,1242}\approx 1.17982\,2547.◄$$

A los polinomios ortonormales que se obtienen al aplicar el proceso de ortogonalización de Gram-Schmidt a la sucesión de polinomios L.I.

$$\{1, x, x^2, x^3, \ldots, x^n, \ldots\}$$

en $C[-1,1]$, se les llaman **polinomios de Legendre** (normalizados). Estos polinomios tienen importantes aplicaciones en la teoría de aproximación, en física, en ecuaciones diferenciales, etc. Los primeros tres polinomios de Legendre normalizados los hemos calculado en el ejemplo precedente. Se puede probar que estos polinomios, antes de normalizar, están dados por la relación

$$w_{k+1}=\frac{k!}{(2k)!}\frac{d^k}{dx^k}\left(x^2-1\right)^k$$

Por ejemplo, para $k=2$,

$$\begin{aligned}w_3 &= \frac{2!}{4!}\frac{d^2}{dx^2}\left(x^2-1\right)^2\\ &= \frac{1}{12}\frac{d}{dx}\left(2\left(x^2-1\right)2x\right)\\ &= \frac{1}{3}\left(3x^2-1\right)\\ &= x^2-\frac{1}{3}\end{aligned}$$

[18]Recuerde que $\|\operatorname{sen}(\pi x)-p_2(x)\|$ fue aproximadamente 0.62615 72471.

que es el mismo polinomio que obtuvimos en el ejemplo anterior. Para $k = 3$, tenemos

$$w_4 = \frac{3!}{6!}\frac{d^3}{dx^3}\left(x^2-1\right)^3$$
$$= x^3 - \frac{3}{5}x,$$

$$\left(\int_{-1}^{1}\left(x^3-\frac{3}{5}x\right)^2 dx\right)^{1/2} = \frac{2}{35}\sqrt{14};$$

de donde

$$u_4 = \frac{1}{\|w_4\|}w_4 = \frac{1}{2}\sqrt{\frac{7}{2}}\left(5x^3-3x\right).$$

De manera similar (los detalles del cálculo se dejan de ejercicio al lector), podemos obtener

$$u_5 = \frac{1}{8}\sqrt{\frac{9}{2}}\left(35x^4-10x^2+3\right) \quad \text{y}$$

$$u_6 = \frac{1}{8}\sqrt{\frac{11}{2}}\left(63x^5-70x^3+15x\right).$$

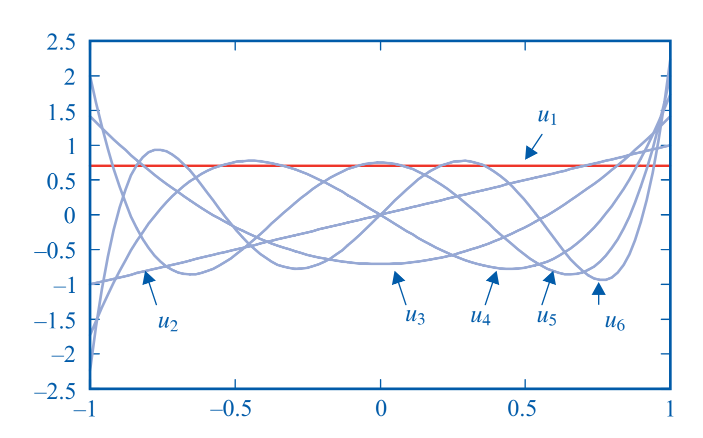


Figura 4-11 • Gráficas de los primeros 6 polinomios de Legendre (normalizados).

En la figura 4-11, se bosquejan las gráficas de los primeros 6 polinomios de Legendre, $u_i(x)$, que calculamos en el ejemplo precedente y en la ulterior discusión.

Es claro que podemos extender este proceso y, dada una función continua f en el espacio $C[-1,1]$ y un número entero no negativo n, construir un polinomio de grado a lo más n, combinación lineal de polinomios de Legendre, que es la mejor aproximación de f, de entre todos los polinomios de grado a lo más n, respecto a la norma inducida por el producto interior en $C[-1,1]$. Así, por ejemplo (se dejan los detalles de los cálculos al lector), la mejor aproximación para $f(x) = \text{sen}(\pi x)$ en $S_4 = \text{gn}(1,x,x^2,x^3,x^4)$, el espacio de polinomios de grado a lo más 4, es

$$p_4 = \frac{35}{2\pi^3}\left(\pi^2-15\right)x^3 + \left(\frac{3}{\pi} - \frac{21}{2\pi^3}\left(\pi^2-15\right)\right)x$$

y la mejor aproximación de f en el espacio de polinomios de grado a lo más 6 es

$$p_6 = \frac{693}{8\pi^5}\left(\pi^4 - 105\pi^2 + 945\right)x^5 + \left(\frac{35}{2\pi^3}\left(\pi^2 - 15\right) - \frac{385}{4\pi^5}\left(\pi^4 - 105\pi^2 + 945\right)\right)x^3$$

$$+ \left(\frac{3}{\pi} - \frac{21}{2\pi^3}\left(\pi^2 - 15\right) + \frac{165}{8\pi^5}\left(\pi^4 - 105\pi^2 + 945\right)\right)x.$$

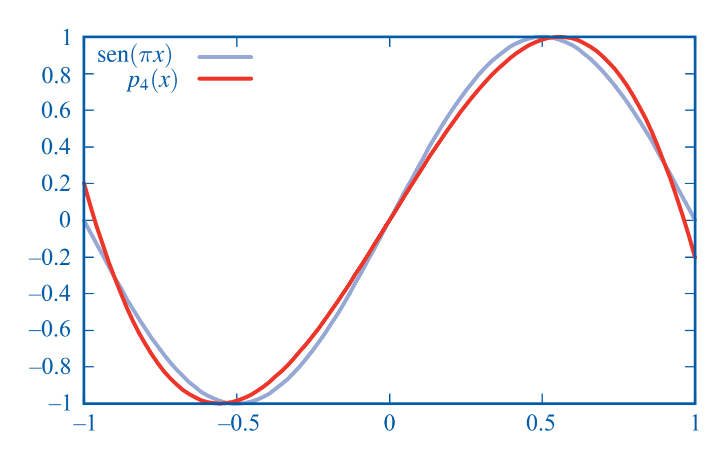


Figura 4-12 • Gráficas de $f(x) = \operatorname{sen}(\pi x)$ y $p_4(x)$. Observe que la diferencia entre las áreas bajo estas curvas (y entre las propias curvas) es muy poca.

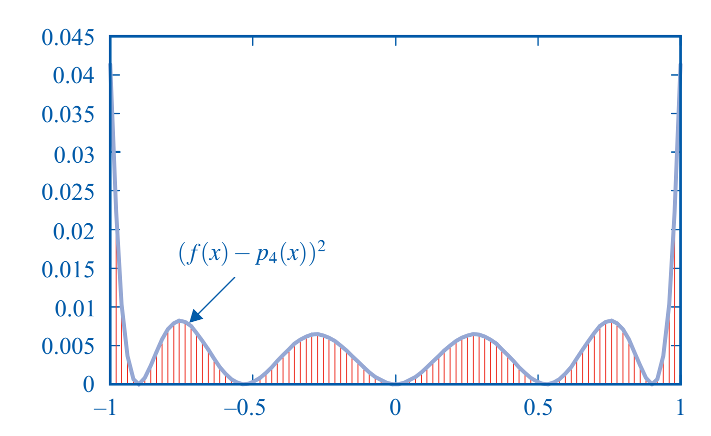


Figura 4-13 • Gráfica de $(f(x) - p_4(x))^2$. El área bajo esta curva es menor que el área bajo la curva $(f(x) - p_2(x))^2$.

Un bosquejo de las gráficas de $f(x) = \operatorname{sen}(\pi x)$ y $p_4(x)$ se ilustra en la figura 4-12. Mientras que la figura 4-13 contiene el área bajo la gráfica de la función $(f(x) - p_4(x))^2$, que ahora es menor que en el caso de aproximación en S_2. Un cálculo un poco laborioso, el cual omitiremos, produce:

$$\|f(x) - p_4(x)\| \approx 9.370289878 \times 10^{-2}$$

y

$$\|f(x) - p_6(x)\| \approx 6.080928379 \times 10^{-3}.$$

Parece que $\|f(x)-p_n(x)\|$ tiende a cero en la medida en que n aumenta. De hecho, esto no es un fenómeno que sólo sucede para $f(x)=\operatorname{sen}(\pi x)$; en realidad, se puede probar que si f es cualquier función continua en $C[-1,1]$ y, para cada $n\in\mathbb{N}$, p_n es el polinomio de grado a lo más n que es la mejor aproximación relativa a la norma inducida por el producto interior para f en el espacio de polinomios de grado[19] a lo más n, entonces

$$\lim_{n\to\infty}\|f-p_n\|=0 \tag{4.25}$$

que equivale a

$$\lim_{n\to\infty}\int_{-1}^{1}(f(x)-p_n(x))^2\,dx=0.$$

Entonces, si se satisface 4.25, la sucesión (p_n) **converge en promedio cuadrático** a la función f. En general, en un espacio vectorial $\mathbf{E}$ con producto interior, se dice que un conjunto ortonormal de vectores $\{\vec{u}_n\}$ es **completo**, si para todo $\vec{v}\in\mathbf{E}$ existe una sucesión $\vec{p}_n$ de combinaciones lineales finitas de los vectores $\vec{u}_i$, que converge en promedio cuadrático a $\vec{v}$. En este sentido, el conjunto de polinomios normalizados de Legendre es completo en $C[-1,1]$.

Aproximación por polinomios trigonométricos

Sea n un entero no negativo y sean las funciones $\psi_0(x)=1$, $\psi_{2k-1}(x)=\cos(kx)$, $\psi_{2k}(x)=\operatorname{sen}(kx)$; $k=1,2,\ldots,n$; esto es, las $2n+1$ funciones continuas $\psi_0(x)=1$, $\psi_1(x)=\cos(x)$, $\psi_2(x)=\operatorname{sen}(x)$, $\psi_3(x)=\cos(2x)$, $\psi_4(x)=\operatorname{sen}(2x),\ldots,\psi_{2n-1}(x)=\cos(nx)$, $\psi_{2n}=\operatorname{sen}(nx)$. Recordemos las identidades trigonométricas

$$\operatorname{sen}(A)\cos(B)=\frac{1}{2}(\operatorname{sen}(A+B)+\operatorname{sen}(A-B)) \tag{4.26}$$

$$\cos(A)\cos(B)=\frac{1}{2}(\cos(A+B)+\cos(A-B)) \tag{4.27}$$

$$\operatorname{sen}(A)\operatorname{sen}(B)=\frac{1}{2}(\cos(A-B)-\cos(A+B)) \tag{4.28}$$

Sean $A,B\in\{0,1,2,\ldots,2n+1\}$, con $A\neq B$. De (4.26):

$$\begin{aligned}\int_0^{2\pi}\operatorname{sen}(Ax)\cos(Bx)dx &= \frac{1}{2}\int_0^{2\pi}(\operatorname{sen}((A+B)x)+\operatorname{sen}((A-B)x))\,dx\\ &= \frac{1}{2}\left[\left(-\frac{1}{A+B}\right)\cos((A+B)x)\Big|_0^{2\pi}\right.\\ &\quad\left.+\left(-\frac{1}{A-B}\right)\cos((A-B)x)\Big|_0^{2\pi}\right]\\ &= 0.\end{aligned}$$

[19] Recuerde que p_n es combinación lineal de polinomios de Legendre normalizados.

De (4.27):

$$\begin{aligned}\int_0^{2\pi} \cos(Ax)\cos(Bx)dx &= \frac{1}{2}\int_0^{2\pi} (\cos((A+B)x)+\cos((A-B)x))\,dx\\ &= \frac{1}{2}\left[\left(\frac{1}{A+B}\right)\operatorname{sen}((A+B)x)\Big|_0^{2\pi}\right.\\ &\qquad \left.+\left(\frac{1}{A-B}\right)\operatorname{sen}((A-B)x)\Big|_0^{2\pi}\right]\\ &= 0.\end{aligned}$$

De (4.28)

$$\begin{aligned}\int_0^{2\pi} \operatorname{sen}(Ax)\operatorname{sen}(Bx)dx &= \frac{1}{2}\int_0^{2\pi} (\cos((A-B)x)-\cos((A+B)x))\,dx\\ &= \frac{1}{2}\left[\left(\frac{1}{A-B}\right)\operatorname{sen}((A-B)x)\Big|_0^{2\pi}\right.\\ &\qquad \left.-\left(\frac{1}{A+B}\right)\operatorname{sen}((A+B)x)\Big|_0^{2\pi}\right]\\ &= 0.\end{aligned}$$

Si $A = B \neq 0$, (4.26) implica

$$\begin{aligned}\int_0^{2\pi} \operatorname{sen}(Ax)\cos(Bx)dx &= \frac{1}{2}\int_0^{2\pi} \operatorname{sen}(2Ax)\,dx\\ &= -\frac{1}{A}\cos(2Ax)|_0^{2\pi}\\ &= 0.\end{aligned}$$

Con lo que hemos probado que

$$\langle \psi_i, \psi_j\rangle = 0 \text{ si } i \neq j$$

y, por tanto, las funciones ψ_j son ortogonales en el espacio $C[0, 2\pi]$. Por otra parte, al utilizar las identidades trigonométricas

$$\operatorname{sen}^2 u = \frac{1-\cos 2u}{2}$$

$$\cos^2 u = \frac{1+\cos 2u}{2}$$

obtenemos

$$\begin{aligned}\int_0^{2\pi} \operatorname{sen}^2(Ax)\,dx &= \frac{1}{2}\int_0^{2\pi} (1-\cos(2Ax))\,dx\\ &= \pi - \frac{1}{4A}\operatorname{sen}(2Ax)|_0^{2\pi}\\ &= \pi\end{aligned}$$

y

$$\begin{aligned}\int_0^{2\pi} \cos^2(Ax)\,dx &= \frac{1}{2}\int_0^{2\pi}(1+\cos(2Ax))\,dx\\ &= \pi + \frac{1}{4A}\operatorname{sen}(2Ax)\Big|_0^{2\pi}\\ &= \pi.\end{aligned}$$

Finalmente,

$$\int_0^{2\pi} dx = 2\pi.$$

Así, el conjunto de funciones $\{\varphi_0, \varphi_1, \ldots, \varphi_{2n}\}$, donde $\varphi_0 = \frac{1}{\sqrt{2\pi}}$, $\varphi_j = \frac{\psi_j}{\sqrt{\pi}}$, $j = 1, 2, \ldots, 2n$; esto es

$$\left\{\begin{array}{c}\varphi_0(x) = \frac{1}{\sqrt{2\pi}}, \varphi_1(x) = \frac{1}{\sqrt{\pi}}\cos(x), \varphi_2(x) = \frac{1}{\sqrt{\pi}}\operatorname{sen}(x),\\ \varphi_3(x) = \frac{1}{\sqrt{\pi}}\cos(2x), \varphi_4(x) = \frac{1}{\sqrt{\pi}}\operatorname{sen}(2x), \ldots\\ \varphi_{2n-1}(x) = \frac{1}{\sqrt{\pi}}\cos(nx), \varphi_{2n}(x) = \frac{1}{\sqrt{\pi}}\operatorname{sen}(nx)\end{array}\right\} \tag{4.29}$$

es un conjunto ortonormal de funciones en $C[0, 2\pi]$. A todo elemento del subespacio de dimensión $2n+1$, $T_n = \operatorname{gn}(\varphi_0, \varphi_1, \ldots, \varphi_{2n}) = \operatorname{gn}(\psi_0, \psi_1, \ldots, \psi_{2n})$, se le llama **polinomio trigonométrico**. Entonces, si f es una función continua en $[0, 2\pi]$, por el teorema 4.11, la proyección de f sobre el espacio de polinomios trigonométricos está dada por

$$p_n = \sum_{k=0}^{2n} \langle f, \varphi_k\rangle \varphi_k$$

donde

$$\langle f, \varphi_k\rangle = \int_0^{2\pi} f(x)\varphi_k(x)dx;$$

que es

$$p_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^{n}(a_k\cos(kx) + b_k\operatorname{sen}(kx)) \tag{4.30}$$

donde, para $k = 0, 1, 2, \ldots, n$,

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(kx)dx \quad \text{y} \quad b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\operatorname{sen}(kx)dx \tag{4.31}$$

A los coeficientes a_k y b_k se les dice **coeficientes de Fourier** de f. El teorema de aproximación garantiza entonces que

$$\min_{q\in S}\|f-q\|$$

se alcanza en p_n. Por lo que p_n es la mejor aproximación para f en el espacio T_n de polinomios trigonométricos para la norma inducida por el producto interior en $C[0,2\pi]$.

▶ **Ejemplo 4.36** Encontrar la mejor aproximación respecto a la norma inducida por el producto interior en $C[0,2\pi]$ de la función $f(x) = e^x$ en el subespacio de polinomios trigonométricos

$$T_2 = \text{gn}(1, \cos x, \text{sen}x, \cos 2x, \text{sen}2x). \blacktriangleleft$$

Solución La mejor aproximación para $f(x) = e^x$ en T_2 está dada por (4.30) para $n = 2$. Esto es:

$$p_2(x) = \frac{1}{2}a_0 + \sum_{k=1}^{2}(a_k \cos(kx) + b_k \text{sen}(kx))$$

donde los coeficientes de Fourier están dados por (4.31). Entonces[20]

$$a_0 = \frac{1}{\pi}\int_0^{2\pi} e^x dx = \frac{e^{2\pi}-1}{\pi}.$$

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_0^{2\pi} e^x \cos(kx)dx \\ &= \frac{1}{\pi}\frac{e^{2\pi}-1}{k^2+1}, \end{aligned}$$

$$\begin{aligned} b_k &= \frac{1}{\pi}\int_0^{2\pi} e^x \text{sen}(kx)dx \\ &= \frac{1}{\pi}\frac{k\left(1-e^{2\pi}\right)}{k^2+1}. \end{aligned}$$

Luego,

$$\begin{aligned} p_2(x) &= \frac{1}{2}\frac{e^{2\pi}-1}{\pi} + \sum_{k=1}^{2}\left[\frac{1}{\pi}\frac{e^{2\pi}-1}{k^2+1}\cos(kx) + \frac{1}{\pi}\frac{k\left(1-e^{2\pi}\right)}{k^2+1}\text{sen}(kx)\right] \\ &= \frac{e^{2\pi}-1}{\pi}\left[\frac{1}{2} + \frac{1}{2}\cos(x) + \frac{1}{5}\cos(2x) - \frac{1}{2}\text{sen}(x) - \frac{2}{5}\text{sen}(2x)\right]. \end{aligned}$$ ✔

Las gráficas de f y su aproximación p_2 en T_2 se encuentran bosquejadas en la figura 4-14; mientras que la gráfica de la función $y = (e^x - p_2(x))^2$ y el área bajo esta curva se muestran en la figura 4-15.

Nuevamente se puede probar que el conjunto de polinomios trigonométricos ortonormales contenidos en (4.29) es completo en $C[0,2\pi]$; es decir, para cualquier función continua f en el intervalo $[0,2\pi]$, existe una sucesión de polinomios trigonométricos p_n, combinación lineal de elementos ortonormales contenidos en el conjunto dado por 4.29, tal que p_n converge en promedio cuadrático a f; esto es,

$$\lim_{n\to\infty} \|f - p_n\| = 0$$

[20] Las integrales $\int e^x \cos(kx)dx$, $\int e^x \text{sen}(kx)dx$, se pueden resolver mediante integración por partes; los detalles de estos cálculos se dejan de ejercicio al lector.

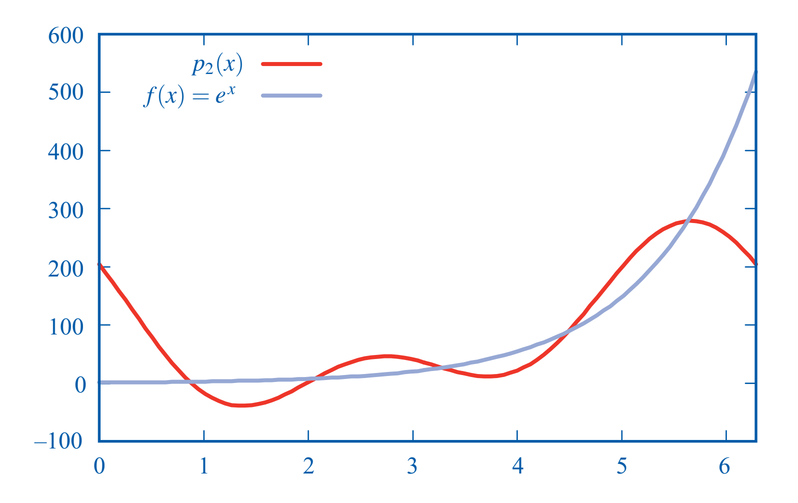


Figura 4-14 • Gráficas de $f(x) = e^x$ y su aproximación $y = p_2(x)$ en T_2.

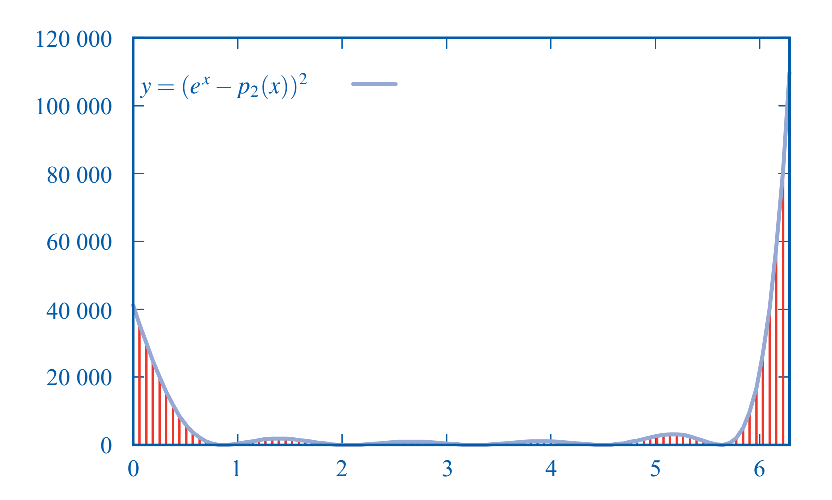


Figura 4-15 • Gráfica de la curva $y = (e^x - p_2(x))^2$ y el área bajo la misma.

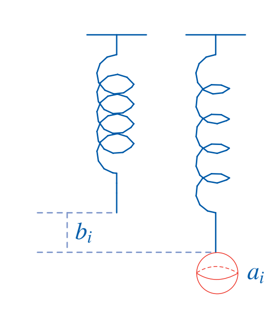
que equivale a

$$\lim_{n\to\infty} \int_0^{2\pi} (f(x) - p_n(x))^2 \, dx = 0,$$

donde p_n está dado por 4.30 y los coeficientes de Fourier a_k, b_k están dados por (4.31).

Mínimos cuadrados

Supongamos que tenemos un conjunto de datos experimentales (a_i, b_i), $i = 1, 2, \ldots, n$. Para fijar ideas, pensemos que hemos colgado de un resorte distintas masas de magnitudes a_i y hemos medido la elongación b_i que producen estos pesos. El propósito es determinar una relación entre el peso que pende y la elongación que produce en el resorte. Para ello, graficamos los pares ordenados (a_i, b_i) obteniendo el bosquejo



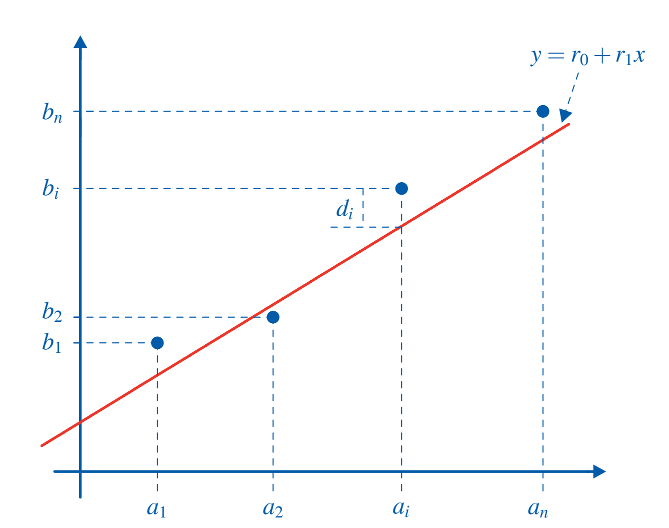


Figura 4-16 • La línea recta $y = r_0 + r_1 x$ que mejor se ajusta al conjunto de datos (a_i, b_i) es, de acuerdo con el criterio de mínimos cuadrados, aquella para la cual $d = \sum_{i=1}^{n} d_i^2$ es mínima, donde $d_i = r_0 + r_1 a_i - b_i$; es decir, la diferencia entre el valor experimental b_i y la ordenada de esta recta en a_i.

contenido en la figura 4-16. Podemos ver en esta figura que parece ser que la relación entre los pares (a_i, b_i) es aproximadamente lineal.[21] Por tanto, lo que se pretende encontrar es la línea recta que mejor se ajuste a estos datos. Entonces, si $y = r_0 + r_1 x$ es la relación buscada, se debe tener $b_i \approx r_0 + r_1 a_i$, $\forall i = 1, 2, \ldots, n$; lo cual se puede expresar matricialmente como

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \approx \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_n \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}$$

O, $\vec{b} \approx A\vec{r}$, con

$$A = \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_n \end{bmatrix}, \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \text{ y } \vec{r} = \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}.$$

Entonces, una manera de definir la recta de mejor ajuste a este conjunto de datos, es pidiendo que sea aquella para la cual la distancia entre $A\vec{r}$ y $\vec{b}$ sea mínima; es decir, eligiendo r_0 y r_1 de tal manera que

$$||\vec{b} - A\vec{r}||$$

sea mínima; lo cual se obtiene por el teorema 4.15, cuando $A\vec{r} = \vec{p}$, con $\vec{p}$ la proyección de $\vec{b}$ sobre el subespacio $W = \text{gn}((1, 1, \ldots, 1, 1), (a_1, a_2, \ldots, a_n))$. Y, entonces, por 4.16

$$\vec{r} = [(A^t A)^{-1} A^t]\vec{b} \tag{4.32}$$

[21] Recuerde que, cuando se recopilan datos experimentales, existen errores en las mediciones.

Notemos, de la figura 4-16, que

$$d = \|\vec{b} - A\vec{r}\|^2 = d_1^2 + d_2^2 + \cdots + d_n^2 \quad (4.33)$$

donde $d_i = r_0 + r_1 a_i - b_i$ para $i = 1, 2, \ldots, n$. Así que, con este criterio, la recta que mejor se ajusta al conjunto de datos (a_i, b_i) es aquella para la cual d es mínima. Dado que se busca el mínimo de la suma de los cuadrados en el lado derecho de 4.33, a este criterio de ajuste se le llama **aproximación** (discreta) **por mínimos cuadrados**.[22]

▶ **Ejemplo 4.37** Ajustar linealmente por mínimos cuadrados el siguiente conjunto de datos:

a_i	2	4	5	6
b_i	6.5	8.5	11	12.5

◀

Solución Para este caso

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix}, \quad A^t = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 5 & 6 \end{bmatrix},$$

$$\begin{aligned} A^t A &= \begin{bmatrix} 4 & 17 \\ 17 & 81 \end{bmatrix}, \\ (A^t A)^{-1} &= \frac{1}{35} \begin{bmatrix} 81 & -17 \\ 17 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 81/35 & -17/35 \\ 17/35 & 4/35 \end{bmatrix}, \\ \vec{r} &= \begin{bmatrix} 81/35 & -17/35 \\ 17/35 & 4/35 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 6.5 \\ 8.5 \\ 11 \\ 12.5 \end{bmatrix} \\ &= \begin{bmatrix} 3.13 \\ 1.53 \end{bmatrix}. \end{aligned}$$

Y la recta de ajuste por mínimos cuadrados de este conjunto de datos es

$$y = 3.13 + 1.53x \quad ✔$$

En general, si se desean ajustar los datos (a_i, b_i) a un polinomio de grado k

$$p_k(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_k x^k,$$

[22] Cuando aproximamos con polinomios de Legendre y polinomios trigonométricos se hizo mínima una integral de la forma $\int_a^b (f(x) - q(x))^2 dx$ para una f dada. A este tipo de optimización también se le dice aproximación (continua) por mínimos cuadrados.

ponemos:

$$\vec{r}=\begin{bmatrix} r_0\\ r_1\\ \vdots\\ r_k\end{bmatrix}\quad \text{y } A=\begin{bmatrix} 1 & a_1 & a_1^2 & \dots & a_1^k\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 1 & a_n & a_n^2 & \dots & a_n^k\end{bmatrix};$$

entonces,

$$\vec{r}=(A^tA)^{-1}A^t\vec{b}.$$

Pues ahora el mínimo de $\|A\vec{r}-\vec{b}\|$ se alcanza cuando $A\vec{r}$ es la proyección de $\vec{b}$ sobre el subespacio

$$W=\text{gn}\big((1,1,\dots,1),\ (a_1,a_2,\dots,a_n),\ (a_1^2,a_2^2,\dots,a_n^2),\ \dots,\ (a_1^k,a_2^k,\dots,a_n^k)\big)$$

de $\mathbb{R}^n$.

▶ **Ejemplo 4.38** Un objeto se deja caer desde un acantilado y se registra su posición, s, desde ese nivel (considerando positiva la dirección hacia abajo); y resultan los datos de la tabla siguiente:

t	1	2	3	4
s	4.85	19.55	44.04	78.35

donde t, el tiempo, se mide en segundos y s en metros. Realizar el ajuste por mínimos cuadrados a un polinomio de grado 2 para estimar $s(t)$ en cualquier instante t. Escriba el resultado final redondeando a una cifra decimal.◀

Solución En este caso:

$$A=\begin{bmatrix} 1 & 1 & 1\\ 1 & 2 & 4\\ 1 & 3 & 9\\ 1 & 4 & 16\end{bmatrix},\quad \vec{b}=\begin{bmatrix} 4.85\\ 19.55\\ 44.04\\ 78.35\end{bmatrix}\quad \text{y}$$

$$\begin{aligned} A^tA &= \begin{bmatrix} 1 & 1 & 1 & 1\\ 1 & 2 & 3 & 4\\ 1 & 4 & 9 & 16\end{bmatrix}\begin{bmatrix} 1 & 1 & 1\\ 1 & 2 & 4\\ 1 & 3 & 9\\ 1 & 4 & 16\end{bmatrix}\\ &= \begin{bmatrix} 4 & 10 & 30\\ 10 & 30 & 100\\ 30 & 100 & 354\end{bmatrix}. \end{aligned}$$

Por el método de Gauss-Jordan, obtenemos (los detalles se dejan de ejercicio al lector):

$$(A^tA)^{-1}=\begin{bmatrix} \frac{31}{4} & -\frac{27}{4} & \frac{5}{4}\\ -\frac{27}{4} & \frac{129}{20} & -\frac{5}{4}\\ \frac{5}{4} & -\frac{5}{4} & \frac{1}{4}\end{bmatrix}.$$

Entonces,

$$\begin{aligned}\vec{r} &= \left[(A^tA)^{-1}A^t\right]\vec{b}\\ &= \begin{bmatrix} \frac{31}{4} & -\frac{27}{4} & \frac{5}{4}\\ -\frac{27}{4} & \frac{129}{20} & -\frac{5}{4}\\ \frac{5}{4} & -\frac{5}{4} & \frac{1}{4}\end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1\\ 1 & 2 & 3 & 4\\ 1 & 4 & 9 & 16\end{bmatrix}\begin{bmatrix} 4.85\\ 19.55\\ 44.04\\ 78.35\end{bmatrix}\\ &= \begin{bmatrix} \frac{9}{4} & -\frac{3}{4} & -\frac{5}{4} & \frac{3}{4}\\ -\frac{31}{20} & \frac{23}{20} & \frac{27}{20} & -\frac{19}{20}\\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4}\end{bmatrix}\begin{bmatrix} 4.85\\ 19.55\\ 44.04\\ 78.35\end{bmatrix}\\ &= \begin{bmatrix} -.0375\\ -.0135\\ 4.9025\end{bmatrix},\end{aligned}$$

que redondeado a una cifra decimal produce

$$\vec{r} = \begin{bmatrix} 0\\ 0\\ 4.9\end{bmatrix}.$$

Luego, el ajuste es

$$y = 4.9t^2. \quad ✔$$

Mínimos cuadrados y factorización QR

Recordemos que, en el método de mínimos cuadrados, al ajustar un conjunto de datos (a_i, b_i) a un polinomio $r_0 + r_1x + \cdots + r_kx^k$, si $\vec{b} = (b_1, \ldots, b_n)$, $\vec{r} = (r_0, r_1, \ldots, r_k)$ y A es la matriz que tiene por columnas a $\begin{bmatrix} 1 & \cdots & 1\end{bmatrix}^t$, $\begin{bmatrix} a_1 & \cdots & a_n\end{bmatrix}^t$, $\begin{bmatrix} a_1^2 & \cdots & a_n^2\end{bmatrix}^t, \ldots, \begin{bmatrix} a_1^k & \cdots & a_n^k\end{bmatrix}^t$, entonces

$$\vec{r} = (A^tA)^{-1}A^t\vec{b}.$$

Sea Q la matriz ortogonal y R la matriz triangular superior del teorema 4.13 para la factorización QR de la matriz A, entonces, puesto que Q es ortogonal ($Q^tQ = I$) y la matriz R es invertible, se tiene

$$\begin{aligned}\vec{r} &= ((QR)^tQR)^{-1}(QR)^t\vec{b}\\ &= (R^tQ^tQR)^{-1}(R^tQ^t)\vec{b}\\ &= (R^tIR)^{-1}(R^tQ^t)\vec{b}\\ &= (R^tR)^{-1}(R^tQ^t)\vec{b}\\ &= R^{-1}(R^t)^{-1}R^tQ^t\vec{b}\\ &= R^{-1}IQ^t\vec{b}\\ &= R^{-1}Q^tb\end{aligned}$$

y, por tanto,

$$R\vec{r} = Q^t\vec{b};$$

luego $\vec{r}$ es la única solución del sistema

$$R\vec{x} = Q^t\vec{b},$$

el cual se puede resolver (puesto que la matriz R es triangular superior) simplemente haciendo sustitución regresiva. Esto puede representar una ventaja en algunos casos al simplificar los cálculos, pues no se tiene que calcular la inversa de la matriz A^tA.

▶ **Ejemplo 4.39** Ajustar, por mínimos cuadrados, el conjunto de datos contenido en la siguiente tabla a un polinomio de grado 3.

a_i	-2	-1	0	1	2
b_i	-3.98	0.90	0	-1.02	3.95

◀

Solución En este caso:

$$A = \begin{bmatrix} 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{bmatrix}$$

y primero aplicamos el proceso de ortogonalización a las columnas de A:

$$\|(1,1,1,1,1)\| = \sqrt{5},$$
$$\vec{u}_1 = \frac{1}{\sqrt{5}}(1,1,1,1,1).$$

$$\begin{aligned}\vec{w}_2 &= (-2,-1,0,1,2) - ((-2,-1,0,1,2)\cdot\vec{u}_1)\vec{u}_1 \\ &= (-2,-1,0,1,2),\end{aligned}$$

$$\vec{u}_2 = \frac{1}{\|\vec{w}_2\|}\vec{w}_2$$
$$\vec{u}_2 = \frac{1}{\sqrt{10}}(-2,-1,0,1,2).$$

$$\begin{aligned}\vec{w}_3 &= (4,1,0,1,4) - ((4,1,0,1,4)\cdot\vec{u}_1)\vec{u}_1 - ((4,1,0,1,4)\cdot\vec{u}_2)\vec{u}_2 \\ &= (4,1,0,1,4) - 2\sqrt{5}\vec{u}_1 - 0\vec{u}_2 \\ &= (2,-1,-2,-1,2),\end{aligned}$$

$$\vec{u}_3 = \frac{1}{\|\vec{w}_3\|}\vec{w}_3$$
$$\vec{u}_3 = \frac{1}{\sqrt{14}}(2,-1,-2,-1,2).$$

$$\begin{aligned}\vec{w}_4 &= (-8,-1,0,1,8) - ((-8,-1,0,1,8)\cdot\vec{u}_1)\vec{u}_1 - ((-8,-1,0,1,8)\cdot\vec{u}_2)\vec{u}_2 \\ &\quad -((-8,-1,0,1,8)\cdot\vec{u}_3)\vec{u}_3 \\ &= (-8,-1,0,1,8) - \vec{0} - \tfrac{17}{5}\sqrt{10}\vec{u}_2 - \vec{0} \\ &= (-8,-1,0,1,8) - \left(-\tfrac{34}{5}, -\tfrac{17}{5}, 0, \tfrac{17}{5}, \tfrac{34}{5}\right) \\ \vec{w}_4 &= \left(-\tfrac{6}{5}, \tfrac{12}{5}, 0, -\tfrac{12}{5}, \tfrac{6}{5}\right),\end{aligned}$$

$$\begin{aligned}\vec{u}_4 &= \tfrac{1}{\|\vec{w}_4\|}\vec{w}_4\\ \vec{u}_4 &= \left(-\tfrac{1}{10}\sqrt{10}, \tfrac{1}{5}\sqrt{10}, 0, -\tfrac{1}{5}\sqrt{10}, \tfrac{1}{10}\sqrt{10}\right).\end{aligned}$$

$$Q = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{10}} & \frac{2}{\sqrt{14}} & -\frac{\sqrt{10}}{10} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{14}} & \frac{\sqrt{10}}{5} \\ \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{14}} & 0 \\ \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{14}} & -\frac{\sqrt{10}}{5} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{10}} & \frac{2}{\sqrt{14}} & \frac{\sqrt{10}}{10} \end{bmatrix}.$$

Y ya que

$$\begin{aligned}
\vec{v}_1 \cdot \vec{u}_1 &= (1,1,1,1,1)\cdot\vec{u}_1\\ &= \sqrt{5},\\
\vec{v}_2 \cdot \vec{u}_1 &= (-2,-1,0,1,2)\cdot\vec{u}_1\\ &= 0,\\
\vec{v}_3 \cdot \vec{u}_1 &= (4,1,0,1,4)\cdot\vec{u}_1\\ &= 2\sqrt{5},\\
\vec{v}_4 \cdot \vec{u}_1 &= (-8,-1,0,1,8)\cdot\vec{u}_1\\ &= 0,\\
\vec{v}_2 \cdot \vec{u}_2 &= (-2,-1,0,1,2)\cdot\vec{u}_2\\ &= \sqrt{10},\\
\vec{v}_3 \cdot \vec{u}_2 &= (4,1,0,1,4)\cdot\vec{u}_2\\ &= 0,\\
\vec{v}_4 \cdot \vec{u}_2 &= (-8,-1,0,1,8)\cdot\vec{u}_2\\ &= \tfrac{17}{5}\sqrt{10},\\
\vec{v}_3 \cdot \vec{u}_3 &= (4,1,0,1,4)\cdot\vec{u}_3\\ &= \sqrt{14},\\
\vec{v}_4 \cdot \vec{u}_3 &= (-8,-1,0,1,8)\cdot\vec{u}_3\\ &= 0,\\
v_4 \cdot \vec{u}_4 &= (-8,-1,0,1,8)\cdot\vec{u}_4\\ &= \tfrac{6}{5}\sqrt{10};
\end{aligned}$$

se tiene que

$$R = \begin{bmatrix} \sqrt{5} & 0 & 2\sqrt{5} & 0 \\ 0 & \sqrt{10} & 0 & \frac{17}{5}\sqrt{10} \\ 0 & 0 & \sqrt{14} & 0 \\ 0 & 0 & 0 & \frac{6}{5}\sqrt{10} \end{bmatrix};$$

y puesto que

$$Q^t\vec{b} = \begin{bmatrix} \frac{1}{5}\sqrt{5} & \frac{1}{5}\sqrt{5} & \frac{1}{5}\sqrt{5} & \frac{1}{5}\sqrt{5} & \frac{1}{5}\sqrt{5} \\ -\frac{1}{5}\sqrt{10} & -\frac{1}{10}\sqrt{10} & 0 & \frac{1}{10}\sqrt{10} & \frac{1}{5}\sqrt{10} \\ \frac{1}{7}\sqrt{14} & -\frac{1}{14}\sqrt{14} & -\frac{1}{7}\sqrt{14} & -\frac{1}{14}\sqrt{14} & \frac{1}{7}\sqrt{14} \\ -\frac{1}{10}\sqrt{10} & \frac{1}{5}\sqrt{10} & 0 & -\frac{1}{5}\sqrt{10} & \frac{1}{10}\sqrt{10} \end{bmatrix} \begin{bmatrix} -3.98 \\ 0.90 \\ 0.00 \\ -1.02 \\ 3.95 \end{bmatrix}$$

$$= \begin{bmatrix} -.03\sqrt{5} \\ 1.394\sqrt{10} \\ 4.2857\times 10^{-3}\sqrt{14} \\ 1.177\sqrt{10} \end{bmatrix}$$

el sistema a resolver es

$$\begin{bmatrix} \sqrt{5} & 0 & 2\sqrt{5} & 0 \\ 0 & \sqrt{10} & 0 & \frac{17}{5}\sqrt{10} \\ 0 & 0 & \sqrt{14} & 0 \\ 0 & 0 & 0 & \frac{6}{5}\sqrt{10} \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} -.03\sqrt{5} \\ 1.394\sqrt{10} \\ 4.2857\times 10^{-3}\sqrt{14} \\ 1.177\sqrt{10} \end{bmatrix}$$

cuya solución es, al hacer sustitución regresiva:

$$\begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} -3.8571\times 10^{-2} \\ -1.9408 \\ 4.2857\times 10^{-3} \\ .98083 \end{bmatrix}.$$

Al redondear a dos cifras decimales, se obtiene el polinomio ajuste de tercer grado

$$p(x) = -0.04 - 1.94x + 0.98x^3.$$

Las gráficas de los datos (a_i, b_i) y el polinomio de ajuste $p(x)$ se encuentran bosquejados en la figura 4-17. ✔

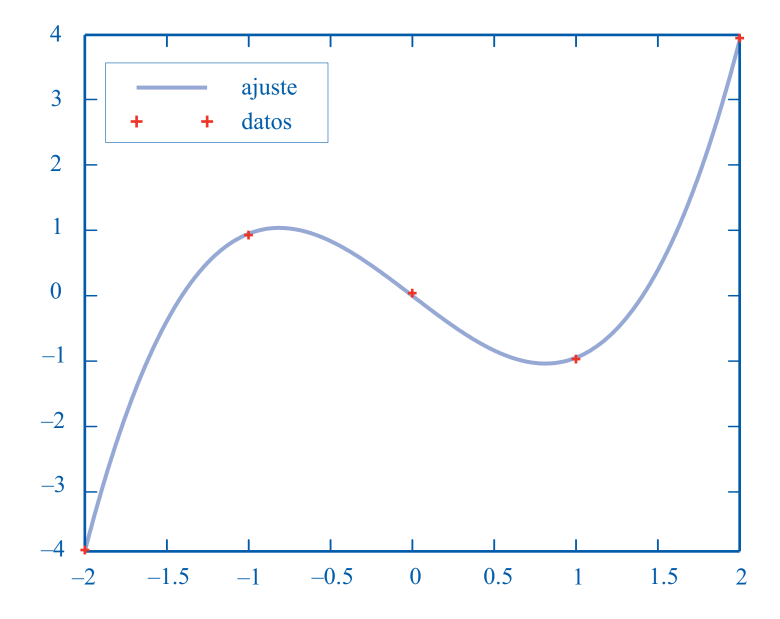


Figura 4-17 •

4.2 Espacios vectoriales normados

En este capítulo hemos visto cómo en un espacio vectorial con producto interior existe una manera natural de precisar la idea de proximidad entre sus vectores por medio de la norma inducida por el producto interior y la distancia que esta norma define. Pero no en todo espacio vectorial se puede definir un producto interior que sea interesante y, sobre todo, útil. Sin embargo, aunque no se cuente con un producto interior en un espacio, la idea de proximidad entre sus elementos se puede llevar a cabo si se generaliza el concepto de norma de un vector; pues, por medio de la norma se puede definir la distancia entre vectores. Para generalizar el concepto de norma (o módulo) en un espacio necesitamos, como hicimos antes en otras generalizaciones (espacio vectorial, producto interior, etc.), abstraer las propiedades más elementales (o esenciales) que tiene la norma en $\mathbb{R}^n$ y en los espacios con producto interior. Estas propiedades son precisamente las que establecimos en los teoremas 3.3 y 4.6 (cfr. páginas 127 y 256).

4.2.1 Definiciones y ejemplos

A partir de aquí, ya no supondremos que los espacios vectoriales con los que se trabaje son espacios con producto interior necesariamente.

Definición 4.11 ***(Norma en un espacio vectorial)*** *Sea* $\mathbf{E}$ *un espacio vectorial. Una norma en* $\mathbf{E}$ *es una función,* $\|\cdot\|:\mathbf{E}\to\mathbb{R}$*, que a cada vector* $\vec{u}\in\mathbf{E}$ *le asigna un número real denotado como* $\|\vec{u}\|$*, llamado la* ***norma*** *(o módulo) de este vector, que satisface las siguientes condiciones:*

1. $\|\vec{u}\|\geq 0 \quad \forall\vec{u}\in\mathbf{E}$.
2. $\|\vec{u}\|=0\Leftrightarrow\vec{u}=\vec{0}_{\mathbf{E}}$.
3. $\|\lambda\vec{u}\|=|\lambda|\,\|\vec{u}\| \quad \forall\lambda\in\mathbb{R},\ \forall\vec{u}\in\mathbf{E}$.
4. $\|\vec{u}+\vec{v}\|\leq\|\vec{u}\|+\|\vec{v}\| \quad \forall\vec{u},\vec{v}\in\mathbf{E}$ (desigualdad triangular).

Si en $\mathbf{E}$ *se ha definido una norma, diremos que* $\mathbf{E}$ *es* un espacio vectorial normado.

► **Ejemplo 4.40** Ya vimos en el teorema 4.6 que si $\mathbf{E}$ es un espacio vectorial con producto interior $\langle\cdot,\cdot\rangle$, entonces $\|\vec{u}\|=\sqrt{\langle\vec{u},\vec{u}\rangle}$ satisface las cuatro condiciones de la definición precedente; por tanto, la norma inducida por el producto interior es una norma en el espacio $\mathbf{E}$. Así, en particular, en los espacios $\mathbb{R}^n$, $\mathfrak{M}_{m\times n}$, ℓ_2 y $C[a,b]$,

$$\|(x_1,x_2,\ldots,x_n)\|=\left(\sum_{i=1}^{n}x_i^2\right)^{1/2},$$

$$\|[a_{ij}]\|=\left(\sum_{\substack{1\leq i\leq m\\1\leq j\leq n}}a_{ij}^2\right)^{1/2}=\operatorname{tra}\left([a_{ij}]^t[a_{ij}]\right),$$

$$\|(x_n)\|=\left(\sum_{n=1}^{\infty}x_n^2\right)^{1/2} \quad \text{y}$$

$$\|f\|=\left(\int_a^b[f(x)]^2dx\right)^{1/2} \quad \text{son normas en sendos espacios.}◄$$

▶ **Ejemplo 4.41** (Norma cúbica) Para cada $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ si definimos[23]

$$\|\vec{u}\|_\infty = \max_{1 \le i \le n} |x_i|,$$

entonces $\|\cdot\|_\infty$ es una norma en $\mathbb{R}^n$, la llamada **norma cúbica**[24] (o norma infinita o norma uniforme). En efecto:

1. Claramente $\|\vec{u}\|_\infty \ge 0 \ \forall \vec{u} \in \mathbb{R}^n$.
2. Es evidente que $\|\vec{0}_{\mathbb{R}^n}\|_\infty = 0$. Supongamos que $\vec{u} = (x_1, x_2, \ldots, x_n)$ y $\|\vec{u}\|_\infty = 0$, entonces $0 \le |x_k| \le \|\vec{u}\|_\infty = 0$ para cada $k = 1, 2, \ldots, n$; por tanto, $|x_k| = 0 \ \forall k$, y por ende $\vec{u} = \vec{0}_{\mathbb{R}^n}$.
3. Si $\lambda \in \mathbb{R}$ y $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, entonces

$$\begin{aligned} \|\lambda\vec{u}\|_\infty &= \max_{1 \le i \le n} |\lambda x_i| \\ &= \max_{1 \le i \le n} |\lambda|\,|x_i| \\ &= |\lambda| \max_{1 \le i \le n} |x_i| \\ &= |\lambda|\,\|\vec{u}\|_\infty . \end{aligned}$$

4. Sean $\vec{u} = (x_1, x_2, \ldots, x_n), \vec{v} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$. Sabemos, por las propiedades del valor absoluto (cfr. pág. 127), que si $a, b \in \mathbb{R}$, entonces $|a+b| \le |a| + |b|$; por tanto,

$$|x_i + y_i| \le |x_i| + |y_i| \quad \forall i = 1, 2, \ldots, n.$$

Dado que, para cada $i = 1, 2, \ldots, n$,

$$\begin{aligned} |x_i| &\le \|\vec{u}\|_\infty \text{ y} \\ |y_i| &\le \|\vec{v}\|_\infty \end{aligned}$$

se tiene

$$|x_i + y_i| \le |x_i| + |y_i| \le \|\vec{u}\|_\infty + \|\vec{v}\|_\infty ;$$

de donde

$$\max_{1 \le i \le n} |x_i + y_i| \le \|\vec{u}\|_\infty + \|\vec{v}\|_\infty .$$

Es decir,

$$\|\vec{u} + \vec{v}\|_\infty \le \|\vec{u}\|_\infty + \|\vec{v}\|_\infty .$$

De 1, 2, 3 y 4, $\|\cdot\|_\infty$ es una norma en $\mathbb{R}^n$. ◀

[23] La notación $\max\limits_{1 \le i \le n} |x_i|$ significa el valor máximo del conjunto de números reales $\{|x_1|, |x_2|, \ldots, |x_n|\}$.

[24] Más adelante aclararemos la razón del nombre de esta norma (norma cúbica) y de la notación $\|\cdot\|_\infty$.

▶ **Ejemplo 4.42** (Norma $\|\cdot\|_1$). Sean $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ y

$$\|\vec{u}\|_1 = |x_1| + |x_2| + \cdots + |x_n| .$$

Mostrar que $\|\cdot\|_1$ es una norma en $\mathbb{R}^n$.◀

DEMOSTRACIÓN ■ 1. Es claro que $\|\vec{u}\|_1 = \sum_{i=1}^n |x_i| \geq 0 \; \forall \vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

2. Es evidente que $\|\vec{0}_{\mathbb{R}^n}\| = 0$. Supongamos que $\vec{u} = (x_1, x_2, \ldots, x_n)$ y que $\|\vec{u}\|_1 = 0$, entonces, para cada $k = 1, 2, \ldots, n$,

$$|x_k| \leq \sum_{i=1}^n |x_i| = \|\vec{u}\|_1 = 0;$$

de donde $|x_k| = 0 \; \forall k = 1, 2, \ldots, n$, por tanto, $\vec{u} = \vec{0}_{\mathbb{R}^n}$.

3. Sean $\lambda \in \mathbb{R}$ y $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, entonces

$$\begin{aligned} \|\lambda\vec{u}\|_1 &= \|(\lambda x_1, \lambda x_1, \ldots, \lambda x_n)\| \\ &= \sum_{i=1}^n |\lambda x_i| \\ &= \sum_{i=1}^n |\lambda|\,|x_i| \\ &= |\lambda| \sum_{i=1}^n |x_i| \\ &= |\lambda|\,\|\vec{u}\|_1 . \end{aligned}$$

4. Sean $\vec{u} = (x_1, x_2, \ldots, x_n), \vec{v} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, entonces

$$\begin{aligned} \|\vec{u} + \vec{v}\|_1 &= \sum_{i=1}^n |x_i + y_i| \\ &\leq \sum_{i=1}^n (|x_i| + |y_i|) \\ &= \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| \\ &= \|\vec{u}\|_1 + \|\vec{v}\|_2 . \end{aligned}$$

De 1, 2, 3 y 4, $\|\cdot\|_1$ es una norma en $\mathbb{R}^n$. ■

▶ **Ejemplo 4.43** (Norma $\|\cdot\|_\infty$ en $\mathfrak{M}_{m\times n}$). Sea $A = [a_{ij}] \in \mathfrak{M}_{m\times n}$, definimos

$$\|A\|_\infty = \max_{\substack{1\leq i\leq m \\ 1\leq j\leq m}} |a_{ij}| .$$

Así, por ejemplo

$$\left\| \begin{bmatrix} -1 & -5 & 3 \\ 2 & 1 & -5 \end{bmatrix} \right\| = 5.$$

De manera análoga al ejemplo 4.41, se puede probar que $\|\cdot\|_\infty$ es una norma en $\mathfrak{M}_{m\times n}$; esta demostración se deja de ejercicio al lector.◀

▶ **Ejemplo 4.44** (Norma $\|\cdot\|_1$ en $\mathfrak{M}_{m\times n}$). En $\mathfrak{M}_{m\times n}$ se define, para cada $A=[a_{ij}]$,

$$\|A\|_1 = \sum_{\substack{1\le i\le m\\1\le j\le n}} |a_{ij}|\,.$$

Análogamente al ejemplo 4.42, se puede demostrar que $\|\cdot\|_1$ es una norma en $\mathfrak{M}_{m\times n}$ (lo cual se deja como ejercicio al lector).◀

❍ **Nota 4.8** Observemos que es posible identificar el espacio de matrices $\mathfrak{M}_{m\times n}$ con el espacio $\mathbb{R}^{mn}$; pues a toda matriz $A=[a_{ij}]\in\mathfrak{M}_{m\times n}$ se le puede asociar el vector

$$\vec{A}=(a_{11},a_{12},\ldots,a_{1n},a_{21},a_{22},\ldots,a_{2n},\ldots,a_{m1},a_{m2},\ldots,a_{mn})\,.$$

Luego,

$$\|A\|_\infty = \|\vec{A}\|_\infty$$

y

$$\|A\|_1 = \|\vec{A}\|_1\,.$$

Por esta identificación, las demostraciones de los dos ejemplos anteriores deben ser evidentes para el lector; y, por esta razón, hemos empleado las mismas notaciones, $\|\cdot\|_\infty$ y $\|\cdot\|_1$, para las normas en $\mathbb{R}^n$ y en $\mathfrak{M}_{m\times n}$.

▶ **Ejemplo 4.45**

1. $\|(-1,5,-7)\|_\infty = \text{máx}\{|-1|,|5|,|7|\}=7$.
2. $\|(-1,5,-7)\|_1 = |-1|+|5|+|-7|=13$.
3. $\left\|\begin{bmatrix}-1 & 2 & 4\\ 1 & -5 & 7\end{bmatrix}\right\|_\infty = \text{máx}\{|-1|,|2|,|4|,|1|,|-5|,|7|\}=7$.
4. $\left\|\begin{bmatrix}-1 & 2 & 4\\ 1 & -5 & 7\end{bmatrix}\right\|_1 = |-1|+|2|+|4|+|1|+|-5|+|7|=20$.◀

▶ **Ejemplo 4.46** (Norma $\|\cdot\|_1$ en $C[a,b]$). Definimos $\|\cdot\|_1 : C[a,b]\to\mathbb{R}$ como

$$\|f\|_1 = \int_a^b |f(x)|dx$$

para cada $f\in C[a,b]$. Mostrar que f es una norma en $C[a,b]$.◀

DEMOSTRACIÓN ■ 1. Dado que la integral de una función no negativa es un número no negativo, se tiene que $\|f\|_1\ge 0$ $\forall f\in C[a,b]$.

2. $\int_a^b 0\cdot dx = 0$, por tanto, $\|\theta\|_1=0$ (recordemos que hemos denotado por θ la función constante cero). Supongamos que $\|f\|_1=0$ y que $f\neq\theta$, entonces existe un punto $x_0\in[a,b]$, tal que $f(x_0)\neq 0$. Por

tanto, $|f(x_0)| > 0$. Por continuidad de $|f|$ existe un intervalo J, de una de estas tres formas: $[a,\delta]$, $[x_0-\delta, x_0+\delta]$ o $[\delta, b]$, contenido en $[a,b]$ tal que $|f(x)| > 0\ \forall x \in J$. Luego,[25]

$$\int_J |f(x)|\,dx > 0.$$

Pero, también

$$\int_J |f(x)|\,dx \le \int_a^b |f(x)|dx = 0;$$

de donde

$$\int_J |f(x)|\,dx = 0.$$

Entonces $\int_J |f(x)|\,dx > 0$ y $\int_J |f(x)|\,dx = 0$; lo cual es una contradicción. Por tanto, el suponer que f no es la función constante cero, aunque $\|f\|_1 = 0$, nos lleva a una conclusión absurda.[26] Así, f tiene que ser necesariamente la función constante cero en $[a,b]$.

3. Si $f \in C[a,b]$ y $\lambda \in \mathbb{R}$, entonces

$$\begin{aligned}\|\lambda f\|_1 &= \int_a^b |\lambda f(x)|\,dx \\ &= \int_a^b |\lambda|\,|f(x)|\,dx \\ &= |\lambda| \int_a^b |f(x)|\,dx \\ &= |\lambda|\,\|f\|_1\,.\end{aligned}$$

4. Si $f, g \in C[a,b]$, entonces

$$\begin{aligned}\|f+g\|_1 &= \int_a^b |f(x)+g(x)|\,dx \\ &\le \int_a^b (|f(x)|+|g(x)|)\,dx \\ &= \int_a^b |f(x)|\,dx + \int_a^b |g(x)|\,dx \\ &= \|f\|_1 + \|g\|_1\,.\end{aligned}$$

De 1, 2, 3 y 4 $\|\cdot\|_1$ es una norma en $C[a,b]$. ■

En la figura 4-18 ilustramos la interpretación geométrica de la norma $\|f\|_1$ en $C[a,b]$.

[25] Hemos utilizado la notación $\int_J |f(x)|\,dx$ para cualquiera de las integrales

$$\int_a^\delta |f(x)|\,dx, \quad \int_{x_0-\delta}^{x_0+\delta} |f(x)|\,dx \quad \text{o} \quad \int_\delta^b |f(x)|\,dx,$$

según se dé el caso.

[26] Si el lector no se siente cómodo con esta demostración, puede intentar utilizar el teorema fundamental del cálculo de manera análoga a como lo hicimos en el ejemplo 4.3 de la página 237.

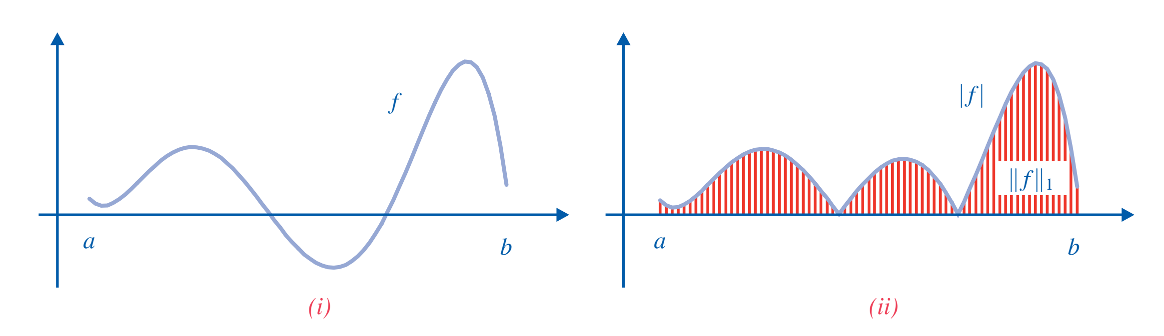


Figura 4-18 • *(i)* Gráfica de una función f en $C[a,b]$. *(ii)* Gráfica de la función $|f|$ y del área bajo esta curva en el intervalo $[a,b]$; $\|f\|_1$ es el valor de esta área.

► **Ejemplo 4.47** (Norma uniforme $\|\cdot\|_\infty$ en $C[a,b]$). Si $f \in C[a,b]$ definimos

$$\|f\|_\infty = \max_{a \le x \le b} |f(x)|,$$

entonces $\|\cdot\|_\infty$ es una norma en $C[a,b]$. A ésta se le llama la **norma uniforme**[27] en el espacio de funciones continuas en el intervalo $[a,b]$. En efecto:

1. $\|f\|_\infty = \max_{a \le x \le b} |f(x)| \ge 0 \quad \forall x \in [a,b]$.
2. Claramente $\|\theta\| = 0$. Puesto que

$$|f(x)| \le \|f\|_\infty \quad \forall x \in [a,b],$$

se sigue que

$$\begin{aligned} \|f\|_\infty = 0 &\Rightarrow |f(x)| \le \|f\|_\infty = 0 \quad \forall x \in [a,b] \\ &\Rightarrow |f(x)| = 0 \quad \forall x \in [a,b] \\ &\Rightarrow f(x) = 0 \quad \forall x \in [a,b] \\ &\Rightarrow f = \theta. \end{aligned}$$

3. Si $f \in C[a,b]$ y $\lambda \in \mathbb{R}$, entonces

$$\begin{aligned} \|\lambda f\|_\infty &= \max_{a \le x \le b} |\lambda f(x)| \\ &= |\lambda| \max_{a \le x \le b} |f(x)| \\ &= |\lambda| \, \|f\|_\infty . \end{aligned}$$

4. Si $f, g \in C[a,b]$, entonces para todo $x \in [a,b]$.

$$\begin{aligned} |f(x) + g(x)| &\le |f(x)| + |g(x)| \\ &\le \|f\|_\infty + \|g\|_\infty . \end{aligned}$$

[27] También se le dice norma cúbica o norma infinito. Hemos utilizado nuevamente la notación $\|\cdot\|_\infty$ por su analogía con esta norma en $\mathbb{R}^n$. La notación $\max_{a \le x \le b} |f(x)|$ significa el máximo de los valores $|f(x)|$ en el intervalo $[a,b]$; la existencia de este valor está garantizada debido a la continuidad de la función en el intervalo cerrado $[a,b]$.

De donde,

$$\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$$

Por 1, 2, 3 y 4 $\|\cdot\|_\infty$ es una norma en $C[a,b]$. ◄

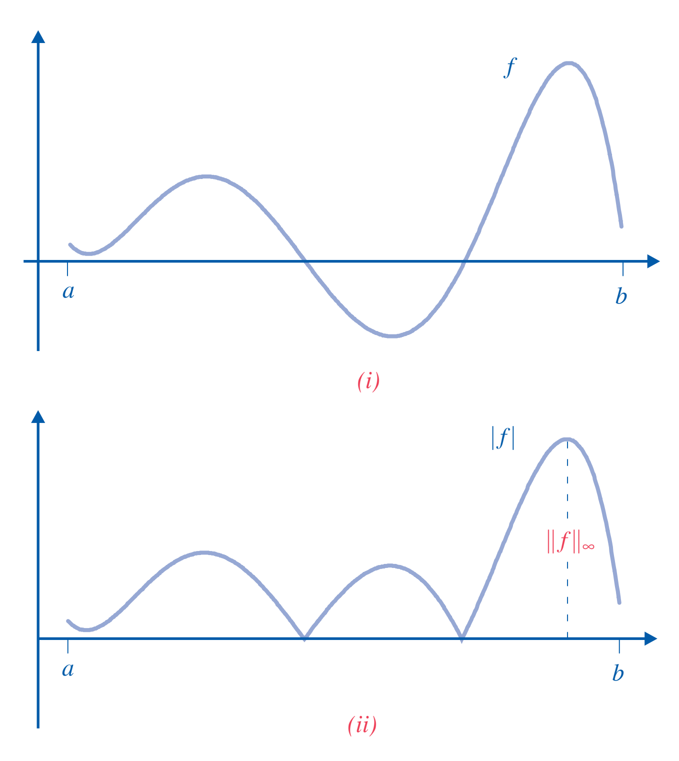


Figura 4-19 • *(i)* Gráfica de una función f en $C[a,b]$. *(ii)* Gráfica de la función $|f|$; la norma $\|f\|_\infty$ es el máximo valor que toma esta función en el intervalo $[a,b]$.

En la figura 4-19, se ilustra el significado geométrico de la norma uniforme en $C[a,b]$.

4.2.2 Distancia en espacios vectoriales normados

Definición 4.12 *Si $\|\cdot\|$ es una norma en un espacio vectorial $\mathbf{E}$ y $\vec{u},\vec{v}\in\mathbf{E}$, se denota y define la distancia entre este par de vectores (distancia relativa a la norma $\|\cdot\|$) como*

$$\mathrm{d}(\vec{u},v) = \|\vec{u}-\vec{v}\|\,.$$

► **Ejemplo 4.48**

1. En $\mathbb{R}^4$ con la norma cúbica,

$$\begin{aligned}\mathrm{d}\left((-1,2,0,1),(2,0,1,1)\right) &= \|(-1,2,0,1)-(2,0,1,1)\|_\infty\\ &= \|(-3,2,-1,0)\|_\infty\\ &= 3.\end{aligned}$$

2. En $\mathbb{R}^4$ con la norma $\|\cdot\|_1$

$$\begin{aligned} \mathrm{d}((-1,2,0,1),(2,0,1,1)) &= \|(-1,2,0,1)-(2,0,1,1)\|_1 \\ &= \|(-3,2,-1,0)\|_1 \\ &= 6. \end{aligned}$$

3. En $\mathfrak{M}_{2\times 2}$ con la norma $\|\cdot\|_\infty$,

$$\begin{aligned} \mathrm{d}\left(\begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\right) &= \left\|\begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\right\|_\infty \\ &= \left\|\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}\right\|_\infty \\ &= 2. \end{aligned}$$

4. En $\mathfrak{M}_{2\times 2}$ con la norma $\|\cdot\|_1$,

$$\begin{aligned} \mathrm{d}\left(\begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\right) &= \left\|\begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\right\|_1 \\ &= \left\|\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}\right\|_1 \\ &= 6. \blacktriangleleft \end{aligned}$$

► **Ejemplo 4.49** En $C[0,2]$ con la norma $\|f\|_1 = \int_0^2 |f(x)|\,dx$, si $f(x) = x^2$ y $f(x) = x$, entonces

$$\begin{aligned} \mathrm{d}(f,g) &= \|f-g\|_1 \\ &= \int_0^2 |f(x)-g(x)|\,dx \\ &= \int_0^2 |x^2-x|\,dx. \end{aligned}$$

Puesto que $x^2 \leq x$ en $[0,1]$ y $x^2 \geq x$ en $[1,2]$, tenemos

$$|f(x)-g(x)| = \begin{cases} x-x^2 & \text{si } 0 \leq x \leq 1, \\ x^2-x & \text{si } 1 \leq x \leq 2; \end{cases}$$

como se ilustra en la figura 4-20. Entonces,

$$\begin{aligned} \int_0^2 |x^2-x|\,dx &= \int_0^1 (x-x^2)\,dx + \int_1^2 (x^2-x)\,dx \\ &= 1. \end{aligned}$$

Luego,

$$\mathrm{d}(f,g) = 1. \blacktriangleleft$$

De manera análoga a la norma inducida por el producto interior $\langle f,g\rangle = \displaystyle\int_a^b f(x)g(x)dx$ en $[a,b]$, la interpretación geométrica de

$$\|f-g\|_1 = \int_a^b |f(x)-g(x)|\,dx$$

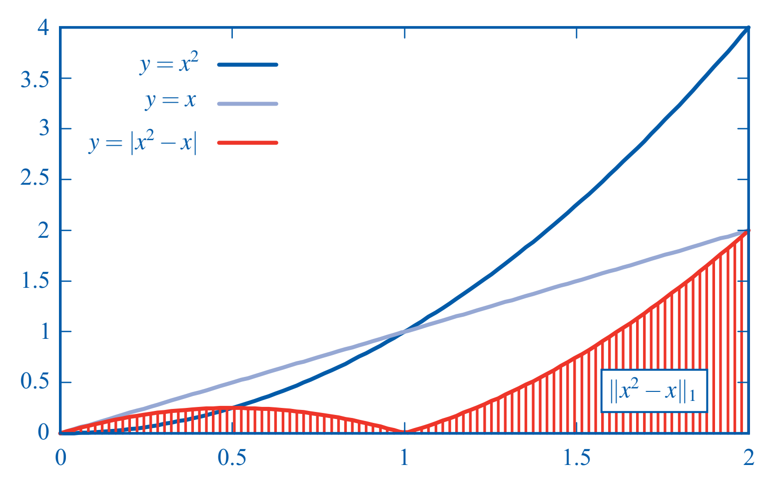


Figura 4-20 • Gráficas de las funciones $y = x^2$; $y = x$; $y = |x^2 - x|$; y el área bajo esta última curva en el intervalo $[0,2]$, $\|x^2 - x\|_1$.

es el valor del área bajo la gráfica de la función $y = |f(x) - g(x)|$ en el intervalo $[a,b]$; como ilustramos en la figura 4-20 para el caso particular de $f(x) = x^2$ y $g(x) = x$.

❍ **Nota 4.9** De aquí en adelante, aunque no se mencione explícitamente, supondremos que $\mathbf{E}$ es un espacio normado con norma $\|\cdot\|$. Los casos concretos, ya sea para espacios o normas, los haremos patentes específicamente.

Bolas y esferas en espacios normados

Definición 4.13 *Si $\vec{u}_0$ es un vector del espacio normado $\mathbf{E}$ y $r > 0$, a los conjuntos*

1. $S(\vec{u}_0, r) = \{\vec{v} \in \mathbf{E} \,|\, \|\vec{u}_0 - \vec{v}\| = r\}$,

2. $B(\vec{u}_0, r) = \{\vec{v} \in \mathbf{E} \,|\, \|\vec{u}_0 - \vec{v}\| < r\}$,

3. $B[\vec{u}_0, r] = \{\vec{v} \in \mathbf{E} \,|\, \|\vec{u}_0 - \vec{v}\| \leq r\}$,

*se les llama **esfera**, **bola abierta** y **bola cerrada**, respectivamente, de centro $\vec{u}_0$ y radio r.*

► **Ejemplo 4.50** En $\mathbb{R}^3$, $\|(x,y,z) - (0,0,0)\| = \sqrt{x^2 + y^2 + z^2}$; luego,

$$S(\vec{0}_{\mathbb{R}^3}, r) = \left\{(x,y,z) \,|\, x^2 + y^2 + z^2 = r^2\right\}$$

Es decir, la esfera de centro $\vec{0}_{\mathbb{R}^3}$ y radio r ilustrada en la figura 4-21.◄

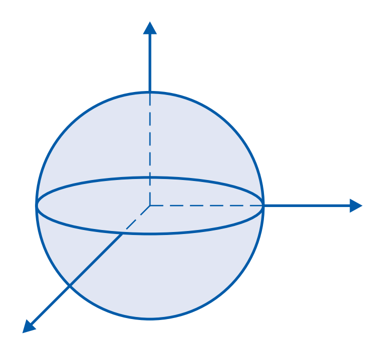


Figura 4-21 • Esfera $S(\vec{0}_{\mathbb{R}^3}, r)$ respecto a la norma $\|(x,y,z)\| = \sqrt{x^2 + y^2 + z^2}$.

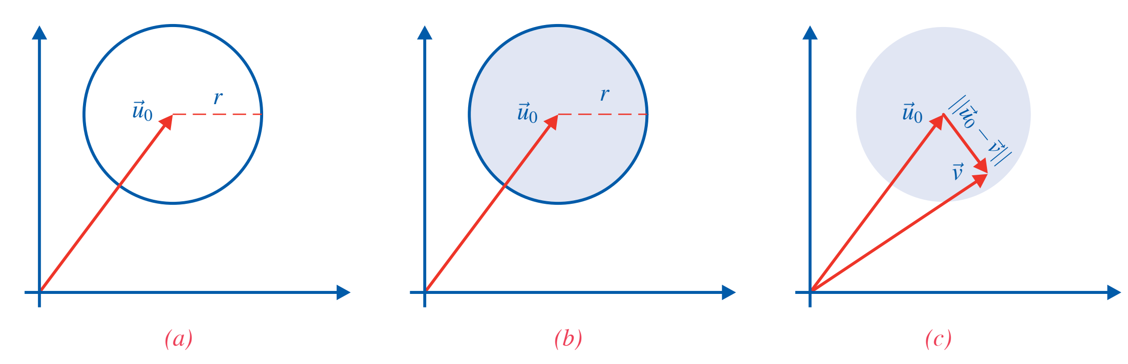


Figura 4-22 • Bolas y esferas en $\mathbb{R}^2$ respecto a la norma $\|(x,y)\| = \sqrt{x^2+y^2}$: *(a)* Esfera de centro $\vec{u}_0$ y radio r; consta de los puntos de la circunferencia de centro $\vec{u}_0$ y radio r. *(b)* Bola cerrada de centro $\vec{u}_0$ y radio r; se compone de todos los puntos de la circunferencia de centro $\vec{u}_0$ y radio r, y los puntos que ésta encierra. *(c)* Bola abierta de centro $\vec{u}_0$ y radio r; está formada por todos los puntos dentro de la circunferencia de centro $\vec{u}_0$ y radio r; i.e., los vectores cuya distancia a $\vec{u}_0$ es inferior a r.

Por generalización, al conjunto del inciso 1 de la definición 4.13 también se le dice esfera. Notemos que el conjunto $B(\vec{0}_{\mathbb{R}^3}, r)$ está formado por todos los puntos dentro de la esfera $S(\vec{0}_{\mathbb{R}^3}, r)$; mientras que $B[\vec{0}_{\mathbb{R}^3}, r]$ contiene a todos los puntos sobre y dentro de la esfera $S(\vec{0}_{\mathbb{R}^3}, r)$. Por analogía a estos conjuntos, es que a $B[\vec{u}_0, r]$ y a $B(\vec{u}_0, r)$ se les llama bola cerrada y bola abierta (de centro $\vec{u}_0$ y radio r), respectivamente, en cualquier espacio normado. Finalmente, notemos que, en cualquier espacio normado, $B(\vec{u}_0, r)$ consiste de los vectores $\vec{v}$ cuya distancia a $\vec{u}_0$ es menor que r; mientras que $B[\vec{u}_0, r]$ está formado por todos los vectores cuya distancia a $\vec{u}_0$ es inferior o igual a r y $S(\vec{u}_0, r)$ consta de los vectores cuya distancia a $\vec{u}_0$ es exactamente r. En la figura 4-22 hemos bosquejado la esfera, la bola abierta y la bola cerrada de centro $\vec{u}_0$ y radio r en $\mathbb{R}^2$ para la norma usual $\|(x,y)\| = \sqrt{x^2+y^2}$.

► **Ejemplo 4.51** Bosquejar

$$B[\vec{0}_{\mathbb{R}^2}, 1] = \{(x,y) \mid \|(x,y)\|_\infty \leq 1\}$$

y

$$B[\vec{0}_{\mathbb{R}^2}, 1] = \{(x,y) \mid \|(x,y)\|_1 \leq 1\}. ◄$$

Solución 1. Si (x,y) pertenece al primer cuadrante, entonces $|x| = x$ y $|y| = y$, luego

$$\begin{array}{ccccccc} 0 & \leq & x & \leq & \|(x,y)\|_\infty & \leq & 1 \quad \text{y} \\ 0 & \leq & y & \leq & \|(x,y)\|_\infty & \leq & 1 \end{array}$$

por tanto,

$$\begin{array}{ccccc} 0 & \leq & x & \leq 1 & \quad \text{y} \\ 0 & \leq & y & \leq 1. & \end{array}$$

Si (x,y) pertenece al segundo cuadrante, entonces $|x| = -x$ y $|y| = y$, así que

$$\begin{array}{ccccccccc} 0 & \leq & -x & = & |x| & \leq & \|(x,y)\|_\infty & \leq & 1 \quad \text{y} \\ 0 & \leq & y & = & |y| & \leq & \|(x,y)\|_\infty & \leq & 1 \end{array}$$

por tanto,

$$\begin{array}{ccccc} -1 & \leq & x & \leq & 0 \quad \text{y} \\ 0 & \leq & y & \leq & 1. \end{array}$$

Si (x,y) pertenece al tercer cuadrante, entonces $|x| = -x$ y $|y| = -y$, así que

$$\begin{array}{ccccccccc} 0 & \leq & -x & = & |x| & \leq & \|(x,y)\|_\infty & \leq & 1 \quad \text{y} \\ 0 & \leq & -y & = & |y| & \leq & \|(x,y)\|_\infty & \leq & 1 \end{array}$$

por tanto,

$$\begin{array}{ccccc} -1 & \leq & x & \leq & 0 \quad \text{y} \\ -1 & \leq & y & \leq & 0. \end{array}$$

Si (x,y) pertenece al cuarto cuadrante, entonces $|x| = x$ y $|y| = -y$, así que

$$\begin{array}{ccccccccc} 0 & \leq & x & = & |x| & \leq & \|(x,y)\|_\infty & \leq & 1 \quad \text{y} \\ 0 & \leq & -y & = & |y| & \leq & \|(x,y)\|_\infty & \leq & 1 \end{array}$$

por lo cual

$$\begin{array}{ccccc} 0 & \leq & x & \leq & 1 \quad \text{y} \\ -1 & \leq & y & \leq & 0. \end{array}$$

Entonces $B[\vec{0}_{\mathbb{R}^2}, 1]$ (relativa a la norma cúbica) es el cuadrado con centro en el origen y arista 2 mostrado en la figura 4-23*(a)*.

2. Si (x,y) pertenece al primer cuadrante, entonces $|x| = x$ y $|y| = y$, entonces

$$x + y = |x| + |y| = \|(x,y)\|_1 \leq 1;$$

por tanto,

$$y \leq 1 - x.$$

Es decir, en este cuadrante los puntos contenidos en $B[\vec{0}_{\mathbb{R}^2}, 1]$ (respecto a la norma $\|\cdot\|_1$) son los puntos sobre y por debajo de la línea recta

$$y = 1 - x.$$

Si (x,y) pertenece al segundo cuadrante, entonces $|x| = -x$ y $|y| = y$, entonces

$$-x + y = |x| + |y| = \|(x,y)\|_1 \leq 1$$

por tanto,

$$y \leq 1 + x.$$

Es decir, en este cuadrante los puntos contenidos en $B[\vec{0}_{\mathbb{R}^2}, 1]$ (respecto a la norma $\|\cdot\|_1$) son los puntos sobre y por debajo de la línea recta

$$y = 1 + x.$$

Si (x,y) pertenece al tercer cuadrante, entonces $|x| = -x$ y $|y| = -y$, luego

$$-x - y = |x| + |y| = \|(x,y)\|_1 \leq 1$$

por tanto,

$$y \geq -1 - x.$$

Es decir, en este cuadrante los puntos contenidos en $B[\vec{0}_{\mathbb{R}^2}, 1]$ (respecto a la norma $\|\cdot\|_1$) son los que están sobre y por encima de la línea recta

$$y = -1 - x.$$

Si (x, y) pertenece al cuarto cuadrante, entonces $|x| = x$ y $|y| = -y$, así

$$x - y = |x| + |y| = \|(x, y)\|_1 \leq 1$$

por tanto,

$$y \geq -1 + x.$$

Es decir, en este cuadrante los puntos contenidos en $B[\vec{0}_{\mathbb{R}^2}, 1]$ (respecto a la norma $\|\cdot\|_1$) son los puntos sobre y por encima de la línea recta

$$y = -1 + x.$$

Entonces, $B[\vec{0}_{\mathbb{R}^2}, 1]$ (relativa a la $\|\cdot\|_1$) es el conjunto de puntos sobre las líneas rectas $y = 1 - x$, $y = 1 + x$, $y = -1 - x$ y $y = -1 + x$; y los puntos dentro del rombo con centro en el origen y lados estas rectas, ilustrado en la figura 4-23*(b)*. ✔

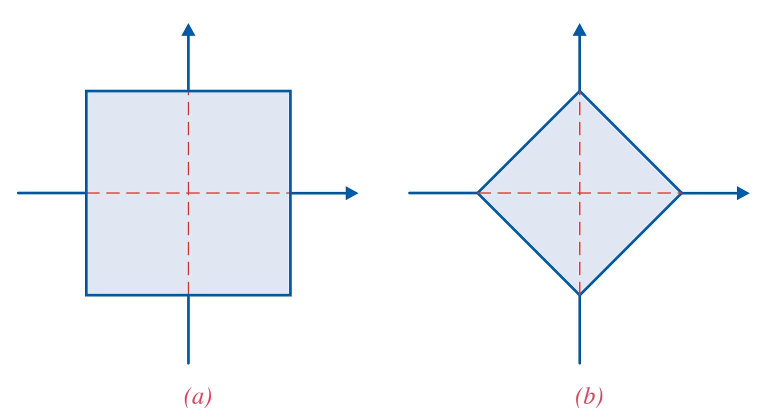


Figura 4-23 • *(a)* $B[\vec{0}_{\mathbb{R}^2}, 1]$, respecto a la norma $\|\cdot\|_\infty$, consta de los puntos sobre y dentro del cuadrilátero de centro el origen y arista 2. *(b)* $B[\vec{0}_{\mathbb{R}^2}, 1]$, respecto a la norma $\|\cdot\|_1$, consiste en los puntos sobre las líneas rectas $y = 1 - x$, $y = 1 + x$, $y = -1 - x$ y $y = -1 + x$ y dentro del rombo con centro el origen y lados estas líneas rectas.

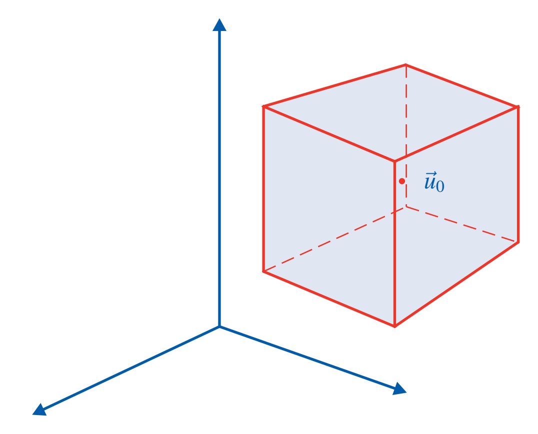


Figura 4-24 • En $\mathbb{R}^3$ la bola cerrada de centro $\vec{u}_0$ y radio r, respecto a la norma $\|\cdot\|_\infty$, es el cubo con centro en este punto y arista $2r$; i.e., los puntos sobre sus seis caras y dentro de la región que éstas encierran.

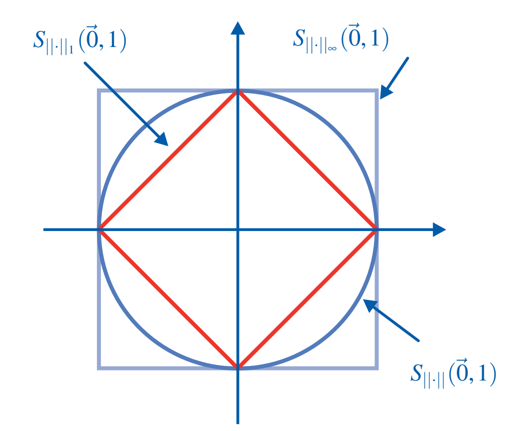


Figura 4-25 • Esferas en $\mathbb{R}^2$ de centro $(0,0)$ y radio 1 relativas a las normas cúbica (cuadrado con centro el origen), la norma canónica $\|(x,y)\| = \sqrt{x^2+y^2}$ (circunferencia con centro en el origen) y la norma $\|\cdot\|_1$ (paralelogramo con centro en el origen).

Es fácil ver, razonando de manera similar al inciso 1 de la solución del ejemplo precedente, que en $\mathbb{R}^3$ geométricamente el conjunto $B[\vec{u}_0, r]$, respecto a la norma $\|\cdot\|_\infty$, es el cubo de centro $\vec{u}_0$ arista $2r$, ilustrado en la figura 4-24. Por esta razón, a la norma $\|\cdot\|_\infty$ se le acostumbra llamar **norma cúbica**. También hemos bosquejado en la figura 4-25 las esferas $S[\vec{0}_{\mathbb{R}^2}, 1]$ para las normas $\|(x,y)\| = \sqrt{x^2+y^2}$, $\|(x,y)\|_\infty$ y $\|(x,y)\|_1$ en $\mathbb{R}^2$ para que el lector pueda compararlas geométricamente.

❍ **Nota 4.10**

1. En los espacios $\mathbb{R}^n$ llamaremos a la norma inducida por el producto punto norma **canónica** (euclidiana, estándar o natural) y la representaremos siempre por el símbolo $\|\cdot\|$, sin subíndices ni supraíndices.
2. En cualquier espacio con producto interior representaremos la norma inducida por este producto con el símbolo $\|\cdot\|$ y, a menos que se diga lo contrario, si no se especifica la norma en este espacio, supondremos que es la norma inducida por el producto interior.
3. Es costumbre llamar a los espacios con producto interior **espacios euclideanos**, debido a su origen intuitivo en los espacios geométricos $\mathbb{R}^2$ y $\mathbb{R}^3$; cuyo estudio sistemático fue llevado a cabo por Euclides, el célebre matemático griego. Por ende, a la norma canónica en $\mathbb{R}^n$ también se le dice norma euclidiana; pues además su definición está inspirada en el teorema de Pitágoras conocido en geometría elemental.
4. Como es natural a las esferas y a las bolas en $\mathbb{R}^2$ se les dice, respectivamente, circunferencias y discos.

► **Ejemplo 4.52** (Bola abierta en $C[a,b]$ respecto a $\|\cdot\|_\infty$). Si $f_0 \in C[a,b]$, bosquejar $B(f_0, r)$ respecto a la norma $\|\cdot\|_\infty$.◄

Solución Si $g \in B(f_0, r)$,

$$\|f_0 - g\|_\infty < r \tag{4.34}$$

Entonces, $\forall x \in [a,b]$

$$|g(x) - f_0(x)| \leq \|f_0 - g\|_\infty < r$$

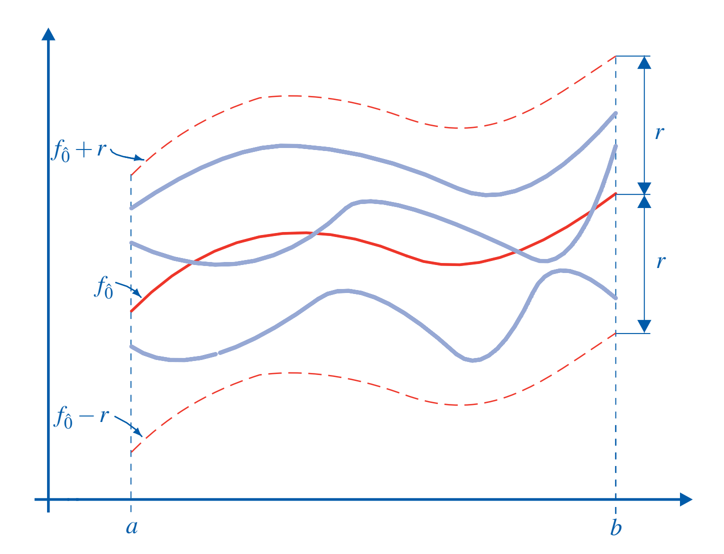


Figura 4-26 • $B(f_0, r)$ en $C[a,b]$, respecto a $\|f\|_\infty$, consta de todas las funciones $g \in C[a,b]$ cuyas gráficas están contenidas entre las funciones $f_0 - r$ y $f_0 + r$, pero que no intersecan a estas funciones.

y, por tanto,

$$-r < g(x) - f_0(x) < r \quad \forall x \in [a,b]\,;$$

lo cual implica

$$f_0(x) - r < g(x) < f_0(x) + r \quad \forall x \in [a,b] \tag{4.35}$$

Inversamente, si se cumple (4.35), se deduce (4.34). Luego $B(f_0, r)$ se compone de todas las funciones continuas en $C[a,b]$, cuya gráfica está contenida entre las gráficas de las funciones $f_0 - r$ y $f_0 + r$; pero que no se intersecan en punto alguno de estas dos gráficas; es decir, las funciones contenidas en la "franja abierta" acotada por las gráficas de las funciones $f_0 - r$ y $f_0 + r$, mostrada en la figura 4-26. ✔

Las bolas, de centro un vector $\vec{u}_0$, son como los intervalos en los números reales; sirven para establecer con precisión el significado de proximidad. Si un intervalo abierto y acotado tiene centro x_0, entonces es de la forma $(x_0 - r, x_0 + r)$ para algún número real $r > 0$; y si $x \in (x_0 - r, x_0 + r)$, $|x - x_0| < r$. Así, la diferencia entre cualquier punto de este intervalo y el centro del mismo, x_0, es inferior a r y este número es el que define el grado de proximidad para puntos respecto a x_0. Por ejemplo, si $r = 10^{-6}$, entonces la diferencia entre x_0 y cualquier punto del intervalo $(x_0 - r, x_0 + r)$ es inferior a 0.000001. Dependiendo del valor de r es el grado de proximidad a x_0 y ese grado de proximidad es relativo y convencional. Lo mismo sucede con las bolas abiertas en espacios normados, pues en estos casos se utiliza simplemente $\|\vec{u}_0 - \vec{u}\|$ en lugar del valor absoluto. Observe que si $g \in B(f_0, r)$ en la figura 4-26 y r es pequeño, dado que $|f_0(x) - g(x)| \le \|f_0 - g\|_\infty < r$, entonces $g(x)$ es muy cercano a $f_0(x)$, para cada $x \in [a,b]$; y, por tanto, la gráfica de la función g es muy próxima a la gráfica de la función f_0; por esta razón, la norma uniforme es una herramienta muy útil para medir proximidad en el espacio de funciones continuas.

❍ **Nota 4.11** Hemos utilizado los símbolos $\|\vec{u}\|_\infty$ y $\|f\|_\infty$ en los espacios $\mathbb{R}^n$ y $C[a,b]$. Cuando se haga uso de este símbolo en $\mathbb{R}^n$, siempre utilizaremos el término *norma cúbica*; mientras que cuando se emplee en $C[a,b]$ diremos que se trata de la *norma uniforme* (o de algunos sinónimos que veremos más adelante). El contexto en cada caso será suficientemente claro para evitar cualquier confusión.

4.2.3 Normas que provienen de productos interiores

Sabemos que todo producto interior $\langle \cdot, \cdot \rangle$ en un espacio vectorial induce una norma en éste, a saber $\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle}$. Supongamos que en el espacio **E** se tiene dada una norma $\|\cdot\|$, surgen naturalmente las siguientes dos cuestiones:

1. ¿Esta norma proviene de un producto interior en **E**? Es decir, ¿existe un producto interior $\langle \cdot, \cdot \rangle$ en **E** cuya norma inducida es precisamente la norma dada $\|\cdot\|$?
2. ¿Bajo qué condiciones se puede definir un producto interior en el espacio **E** por medio de esa norma, de tal suerte que coincida con la norma inducida por ese producto interior?

Por fortuna, estos dos interrogantes tienen una respuesta completa; para dar ésta haremos uso nuevamente del origen geométrico que tienen el producto interior y la norma.

Identidad del paralelogramo

Consideremos el paralelogramo ilustrado en *(i)*, *(ii)* y *(iii)* de la figura 4-27:

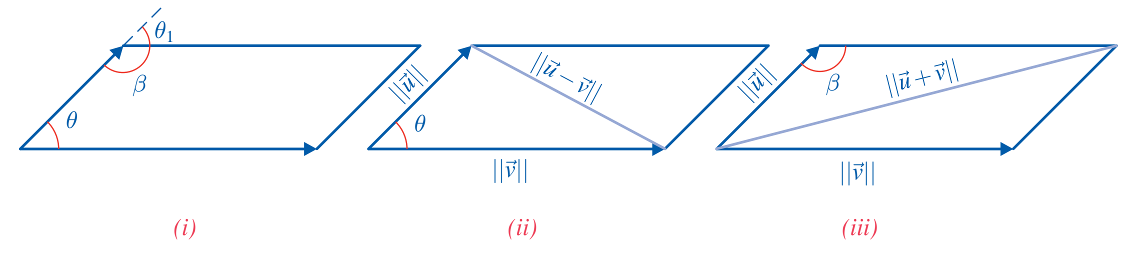


Figura 4-27 •

- De *(i)* de esta figura se desprende que $\theta = \theta_1$ por ser ángulos correspondientes y, por tanto,

$$\beta = 180 - \theta .$$

- De *(ii)* y la ley de cosenos, tenemos

$$\|\vec{u}-\vec{v}\|^2 = \|u\|^2 + \|\vec{v}\|^2 - 2\,\|\vec{u}\|\,\|\vec{v}\| \cos\theta \tag{4.36}$$

- De *(iii)* y la ley de cosenos, se desprende

$$\begin{aligned} \|\vec{u}+\vec{v}\|^2 &= \|u\|^2 + \|\vec{v}\|^2 - 2\,\|\vec{u}\|\,\|\vec{v}\| \cos\beta \\ &= \|u\|^2 + \|\vec{v}\|^2 - 2\,\|\vec{u}\|\,\|\vec{v}\| \cos(180-\theta) \\ &= \|u\|^2 + \|\vec{v}\|^2 - 2\,\|\vec{u}\|\,\|\vec{v}\| (-\cos(\theta)) \quad \therefore \end{aligned}$$

$$\|\vec{u}+\vec{v}\|^2 = \|u\|^2 + \|\vec{v}\|^2 + 2\,\|\vec{u}\|\,\|\vec{v}\| \cos(\theta) \tag{4.37}$$

- Tenemos entonces de (4.36) y (4.37)

$$\|\vec{u}+\vec{v}\|^2 + \|\vec{u}-\vec{v}\|^2 = 2\left(\|u\|^2 + \|\vec{v}\|^2\right) \tag{4.38}$$

La igualdad (4.38) se llama **identidad del paralelogramo** y es un hecho conocido de geometría elemental: *en todo paralelogramo la suma de los cuadrados de las longitudes de sus diagonales es igual a la suma de los cuadrados de las longitudes de los lados del paralelogramo.*

Ahora supongamos que tenemos un espacio vectorial $\mathbf{E}$ con producto interior $\langle\cdot,\cdot\rangle$. Sean $\vec{u},\vec{v}$ un par de vectores en este espacio, entonces

$$\begin{aligned} \|\vec{u}+\vec{v}\|^2 &= \langle\vec{u}+\vec{v},\vec{u}+\vec{v}\rangle \\ &= \langle\vec{u},\vec{u}\rangle+\langle\vec{u},\vec{v}\rangle+\langle\vec{v},\vec{u}\rangle+\langle\vec{v},\vec{v}\rangle \quad \therefore \\ \|\vec{u}+\vec{v}\|^2 &= \|\vec{u}\|^2+2\langle\vec{u},\vec{v}\rangle+\|\vec{v}\|^2 \end{aligned} \tag{4.39}$$

y

$$\begin{aligned} \|\vec{u}-\vec{v}\|^2 &= \langle\vec{u}-\vec{v},\vec{u}-\vec{v}\rangle \\ &= \langle\vec{u},\vec{u}\rangle-\langle\vec{u},\vec{v}\rangle-\langle\vec{v},\vec{u}\rangle+\langle\vec{v},\vec{v}\rangle \quad \therefore \\ \|\vec{u}-\vec{v}\|^2 &= \|\vec{u}\|^2-2\langle\vec{u},\vec{v}\rangle+\|\vec{v}\|^2 \end{aligned} \tag{4.40}$$

Al sumar lado a lado las igualdades (4.39) y (4.40) se obtiene la identidad del paralelogramo en espacios con producto interior:

$$\|\vec{u}+\vec{v}\|^2+\|\vec{u}-\vec{v}\|^2=2\left(\|\vec{u}\|^2+\|\vec{v}\|^2\right)$$

Hemos probado así el siguiente teorema.

Teorema 4.16 ***(Identidad del paralelogramo)*** *Sea* $\mathbf{E}$ *un espacio con producto interior* $\langle\cdot,\cdot\rangle$ *y norma inducida* $\|\cdot\|$*. Entonces,* $\forall\vec{u},\vec{v}\in\mathbf{E}$*, se cumple*

$$\|\vec{u}+\vec{v}\|^2+\|\vec{u}-\vec{v}\|^2=2\left(\|\vec{u}\|^2+\|\vec{v}\|^2\right) \tag{4.41}$$

Por otra parte, si $\mathbf{E}$ es un espacio normado con norma $\|\cdot\|$ que proviene de un producto interior, entonces la norma dada y la inducida por este producto deben coincidir y, por tanto, $\|\cdot\|$ tiene que satisfacer la identidad del paralelogramo (4.41) del teorema 4.16. Es decir, para que una norma provenga de un producto interior es necesario que satisfaga la identidad del paralelogramo. Resumimos este resultado en el siguiente teorema.

Teorema 4.17 *Sea* $\mathbf{E}$ *un espacio normado con norma* $\|\cdot\|$*. Para que* $\|\cdot\|$ *provenga de un producto interior, es necesario que esta norma cumpla con la identidad del paralelogramo (4.41).*

Así, para probar que una norma no proviene de algún producto interior basta probar, con un contraejemplo, que no cumple con la identidad del paralelogramo.

▶ **Ejemplo 4.53** Determinar si la norma cúbica, $\|\cdot\|_\infty$, proviene de un producto interior en $\mathbb{R}^n$.◀

Solución Sean $\vec{u}=(1,2,0,\ldots,0), \vec{v}=(2,-3,0,\ldots,0)\in\mathbb{R}^n$. Por un lado

$$\begin{aligned}\|\vec{u}+\vec{v}\|_\infty^2+\|\vec{u}-\vec{v}\|_\infty^2 &= \|(3,-1,0\ldots,0)\|_\infty^2+\|(-1,5,0\ldots,0)\|_\infty^2\\ &= 9+25\\ &= 34\end{aligned}$$

y, por otro,

$$\begin{aligned}2\left(\|\vec{u}\|_\infty^2+\|\vec{v}\|_\infty^2\right) &= 2\left(\|(1,2,0,\ldots,0)\|_\infty^2+\|(2,-3,0,\ldots,0)\|_\infty^2\right)\\ &= 2\,(4+9)\\ &= 26.\end{aligned}$$

De donde se desprende que la norma cúbica no cumple con la identidad del paralelogramo (4.41), por lo que no proviene de un producto interior. ✔

❍ **Nota 4.12** El lector debe reflexionar con cuidado el significado del hecho mostrado en el ejemplo anterior: *no existe producto interior en* $\mathbb{R}^n$ *cuya norma inducida sea la norma cúbica.*

Supongamos nuevamente que tenemos un espacio con producto interior $\langle\cdot,\cdot\rangle$ y norma inducida por este producto $\|\cdot\|$, entonces, al restar miembro a miembro las igualdades (4.39) y (4.40) se obtiene

$$\|\vec{u}+\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2=4\,\langle\vec{u},\vec{v}\rangle\,;$$

de donde,

$$\langle\vec{u},\vec{v}\rangle=\frac{1}{4}\left[\|\vec{u}+\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2\right].$$

Con lo que hemos probado el siguiente teorema.

Teorema 4.18 ***(Identidad de polarización)*** *Sea* **E** *un espacio con producto interior* $\langle\cdot,\cdot\rangle$ *y norma inducida* $\|\cdot\|$. *Entonces,* $\forall\vec{u},\vec{v}\in\mathbf{E}$ *se cumple la siguiente igualdad*

$$\langle\vec{u},\vec{v}\rangle=\frac{1}{4}\left[\|\vec{u}+\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2\right] \tag{4.42}$$

llamada ***identidad de polarización****.*

Así, para que una norma $\|\cdot\|$ provenga de un producto interior $\langle\cdot,\cdot\rangle$, necesita primero satisfacer la identidad del paralelogramo (4.41); y en caso de que así sea, el producto interior de donde proviene debe cumplir con la identidad de polarización (4.42). Entonces, conjeturamos que si una norma satisface la identidad del paralelogramo, el producto interior definido por (4.42) es, efectivamente, un producto escalar cuya norma inducida es precisamente esta norma. En el siguiente teorema probamos esta conjetura.

Teorema 4.19 *Sea* **E** *un espacio normado con norma* $\|\cdot\|$ *que satisface la identidad del paralelogramo (4.41). Si se define, para cada* $\vec{u},\vec{v}\in\mathbf{E}$,

$$\langle\vec{u},\vec{v}\rangle=\frac{1}{4}\left[\|\vec{u}+\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2\right],$$

entonces $\langle\cdot,\cdot\rangle$ *es un producto interior cuya norma inducida es precisamente* $\|\cdot\|$.

DEMOSTRACIÓN 1. Sea $\vec{u}\in\mathbf{E}$, entonces

$$\begin{aligned}\langle\vec{u},\vec{u}\rangle&=\frac{1}{4}\left[\|\vec{u}+\vec{u}\|^2-\|\vec{0}_{\mathbf{E}}\|^2\right]\\&=\frac{1}{4}\|2\vec{u}\|^2\\&=\frac{4}{4}\|\vec{u}\|^2\\&=\|\vec{u}\|^2\geq 0\end{aligned}$$

y, por tanto,

$$\langle\vec{u},\vec{u}\rangle=0\Leftrightarrow\vec{u}=\vec{0}_{\mathbf{E}}.$$

Puesto que

$$\sqrt{\langle\vec{u},\vec{u}\rangle}=\|\vec{u}\|,$$

siendo $\|\cdot\|$ la norma dada del espacio, se sigue que, de ser $\langle\cdot,\cdot\rangle$ un producto interior, la norma inducida coincide con la norma dada en el espacio; es decir, $\|\cdot\|$ proviene de este producto interior.

2. Si $\vec{u},\vec{v}\in\mathbf{E}$,

$$\begin{aligned}\langle\vec{u},\vec{v}\rangle&=\frac{1}{4}\left[\|\vec{u}+\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2\right]\\&=\frac{1}{4}\left[\|\vec{v}+\vec{u}\|^2-\|\vec{v}-\vec{u}\|^2\right]\\&=\langle\vec{v},\vec{u}\rangle.\end{aligned}$$

3. Sean $\vec{u},\vec{v}\in\mathbf{E}$, entonces

$$\langle\vec{u},\vec{v}\rangle=\frac{1}{4}\left[\|\vec{u}+\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2\right].$$

Por la identidad del paralelogramo (4.41) (recuerde que la norma de este espacio la cumple por hipótesis)

$$\|\vec{u}+\vec{v}\|^2=2\|\vec{u}\|^2+2\|\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2.$$

Y, por tanto,

$$\langle\vec{u},\vec{v}\rangle=\frac{1}{4}\left[2\|\vec{u}\|^2+2\|\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2\right]$$

esto es,

$$\langle\vec{u},\vec{v}\rangle = \frac{1}{2}\left[\|\vec{u}\|^2 + \|\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2\right] \tag{4.43}$$

Sean ahora $\vec{u},\vec{v},\vec{w} \in \mathbf{E}$ tres vectores arbitrarios. Entonces, por (4.43),

$$\langle\vec{u},\vec{v}\rangle = \frac{1}{2}\left[\|\vec{u}\|^2 + \|\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2\right]$$

y

$$\langle\vec{u},\vec{w}\rangle = \frac{1}{2}\left[\|\vec{u}\|^2 + \|\vec{w}\|^2 - \|\vec{u}-\vec{w}\|^2\right].$$

Donde

$$\langle\vec{u},\vec{v}\rangle + \langle\vec{u},\vec{w}\rangle = \frac{1}{2}\left[\|\vec{u}\|^2 + \|\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2 + \|\vec{u}\|^2 + \|\vec{w}\|^2 - \|\vec{u}-\vec{w}\|^2\right].$$

Por otra parte,

$$\langle\vec{u},\vec{v}+\vec{w}\rangle = \frac{1}{4}\left[\|\vec{u}+\vec{v}+\vec{w}\|^2 - \|\vec{u}-\vec{v}-\vec{w}\|^2\right].$$

Por la identidad del paralelogramo

$$\begin{aligned}\|\vec{u}+\vec{v}+\vec{w}\|^2 &= \|(\vec{u}+\vec{v})+\vec{w}\|^2\\ &= 2\|\vec{u}+\vec{v}\|^2 + 2\|\vec{w}\|^2 - \|\vec{u}+\vec{v}-\vec{w}\|^2\\ &= 2\left(2\|\vec{u}\|^2 + 2\|\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2\right) + 2\|\vec{w}\|^2\\ &\quad - \|\vec{u}+\vec{v}-\vec{w}\|^2\end{aligned}$$

y

$$\begin{aligned}\|\vec{u}-\vec{v}-\vec{w}\|^2 &= \|(\vec{u}-\vec{w})-\vec{v}\|^2\\ &= 2\|\vec{u}-\vec{w}\|^2 + 2\|\vec{v}\|^2 - \|(\vec{u}-\vec{w})+\vec{v}\|^2.\end{aligned}$$

Por tanto,

$$\begin{aligned}\langle\vec{u},\vec{v}+\vec{w}\rangle &= \frac{1}{4}\Big[2\left(2\|\vec{u}\|^2 + 2\|\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2\right) + 2\|\vec{w}\|^2\\ &\quad - \|\vec{u}+\vec{v}-\vec{w}\|^2 - 2\|\vec{u}-\vec{w}\|^2 - 2\|\vec{v}\|^2\\ &\quad + \|(\vec{u}-\vec{w})+\vec{v}\|^2\Big]\\ &= \frac{1}{4}\left[4\|\vec{u}\|^2 + 2\|\vec{v}\|^2 - 2\|\vec{u}-\vec{v}\|^2 + 2\|\vec{w}\|^2 - 2\|\vec{u}-\vec{w}\|^2\right]\\ &= \frac{1}{4}\Big[2\|\vec{u}\|^2 + 2\|\vec{v}\|^2 - 2\|\vec{u}-\vec{v}\|^2 + 2\|\vec{u}\|^2 + 2\|\vec{w}\|^2\\ &\quad -2\|\vec{u}-\vec{w}\|^2\Big]\\ &= \frac{1}{2}\left[\|\vec{u}\|^2 + \|\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2 + \|\vec{u}\|^2 + \|\vec{w}\|^2 - \|\vec{u}-\vec{w}\|^2\right]\end{aligned}$$

Esto es,

$$\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle \tag{4.44}$$

4. Sean $\vec{u}, \vec{v} \in \mathbf{E}$ un par de vectores fijos pero arbitrarios. Mostraremos por inducción[28] que para todo entero no negativo m se cumple

$$\langle \vec{u}, m\vec{v} \rangle = m \langle \vec{u}, \vec{v} \rangle \tag{4.45}$$

Para $m = 1$,

$$\langle \vec{u}, \vec{v} \rangle = 1 \langle \vec{u}, \vec{v} \rangle .$$

Sea $k > 1$ un entero y supongamos que:

$$\langle \vec{u}, k\vec{v} \rangle = k \langle \vec{u}, \vec{v} \rangle \tag{4.46}$$

Entonces,

$$\langle \vec{u}, (k+1)\vec{v} \rangle = \langle \vec{u}, k\vec{v} + \vec{v} \rangle$$

y por (4.44) y la hipótesis de inducción (4.46)

$$\begin{aligned} \langle \vec{u}, (k+1)\vec{v} \rangle &= \langle \vec{u}, k\vec{v} + \vec{v} \rangle \\ &= \langle \vec{u}, k\vec{v} \rangle + \langle \vec{u}, \vec{v} \rangle \\ &= k \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{v} \rangle \\ &= (k+1) \langle \vec{u}, \vec{v} \rangle . \end{aligned}$$

Por inducción se sigue que (4.45) es válida para todo entero no negativo m. Por otra parte,

$$\begin{aligned} \langle \vec{u}, -\vec{v} \rangle &= \frac{1}{4} \left[\|\vec{u} - \vec{v}\|^2 - \|\vec{u} + \vec{v}\|^2 \right] \\ &= -\frac{1}{4} \left[\|\vec{u} + \vec{v}\|^2 - \|\vec{u} - \vec{v}\|^2 \right] \\ &= -\langle \vec{u}, \vec{v} \rangle \end{aligned}$$

Por tanto,

$$\langle \vec{u}, m\vec{v} \rangle = m \langle \vec{u}, \vec{v} \rangle$$

para todo entero m. Sea ahora $\lambda \neq 0$ un número real. Entonces, por (4.43),

$$\begin{aligned} \langle \lambda\vec{u}, \vec{v} \rangle &= \frac{1}{2} \left[\|\lambda\vec{u}\|^2 + \|\vec{v}\|^2 - \|\lambda\vec{u} - \vec{v}\|^2 \right] \\ &= \frac{1}{2} \left[\lambda^2 \|\vec{u}\|^2 + \lambda^2 \|(1/\lambda)\vec{v}\|^2 - \lambda^2 \|\vec{u} - (1/\lambda)\vec{v}\|^2 \right] \\ &= \lambda^2 \frac{1}{2} \left[\|\vec{u}\|^2 + \|(1/\lambda)\vec{v}\|^2 - \|\vec{u} - (1/\lambda)\vec{v}\|^2 \right] \\ &= \lambda^2 \langle \vec{u}, (1/\lambda)\vec{v} \rangle . \end{aligned}$$

[28] En el apéndice A se puede consultar un breve estudio del principio de inducción, herramienta fundamental para demostraciones que involucran a los números enteros.

Sea $q \neq 0$ un entero, entonces

$$\begin{aligned}\left\langle \frac{1}{q}\vec{u}, \vec{v}\right\rangle &= \frac{1}{q^2}\langle \vec{u}, q\vec{v}\rangle \\ &= \frac{q}{q^2}\langle \vec{u}, \vec{v}\rangle \\ &= \frac{1}{q}\langle \vec{u}, \vec{v}\rangle.\end{aligned}$$

En síntesis, hemos probado que

$$\langle \vec{u}, m\vec{v}\rangle = m\langle \vec{u}, \vec{v}\rangle$$

para todo entero m,

$$\left\langle \frac{1}{q}\vec{u}, \vec{v}\right\rangle = \frac{1}{q}\langle \vec{u}, \vec{v}\rangle$$

para todo entero $q \neq 0$; y por simetría se tiene también

$$\begin{aligned}\langle m\vec{u}, \vec{v}\rangle &= m\langle \vec{u}, \vec{v}\rangle, \\ \left\langle \vec{u}, \frac{1}{q}\vec{v}\right\rangle &= \frac{1}{q}\langle \vec{u}, \vec{v}\rangle.\end{aligned}$$

Sean ahora p, q un par de enteros con $q \neq 0$. Entonces,

$$\begin{aligned}\left\langle \frac{p}{q}\vec{u}, \vec{v}\right\rangle &= \left\langle p\left(\frac{1}{q}\vec{u}\right), \vec{v}\right\rangle \\ &= p\left\langle \frac{1}{q}\vec{u}, \vec{v}\right\rangle \\ &= \frac{p}{q}\langle \vec{u}, \vec{v}\rangle.\end{aligned}$$

Sea la función $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(\lambda) = \|\lambda\vec{u}\|.$$

Entonces, por la desigualdad triangular,

$$\begin{aligned}|f(\lambda_1) - f(\lambda_2)| &= |\|\lambda_1\vec{u}\| - \|\lambda_2\vec{u}\|| \\ &\leq \|\lambda_1\vec{u} - \lambda_2\vec{u}\| \\ &= |\lambda_1 - \lambda_2|\,\|\vec{u}\|;\end{aligned}$$

de donde f es una función continua en $\mathbb{R}$ y, por tanto, la función

$$\varphi(\lambda) = \frac{1}{2}\left[\|\lambda\vec{u}\|^2 + \|\vec{v}\|^2 - \|\lambda\vec{u} - v\|^2\right]$$

también es continua en $\mathbb{R}$. Sea $\lambda \in \mathbb{R}$ un número real cualquiera fijo; sea p_n una sucesión de números racionales $p_n = a_n/b_n$ que converge a λ; esto es,

$$\lim_{n\to\infty} p_n = \lambda.$$

Por la continuidad de[29] φ,

$$\lim_{n\to\infty} \varphi(p_n) = \varphi(\lambda);$$

es decir,

$$\begin{aligned}\langle \lambda\vec{u},\vec{v}\rangle &= \varphi(\lambda)\\ &= \lim_{n\to\infty}\varphi(p_n)\\ &= \lim_{n\to\infty}\frac{1}{2}\left[\|p_n\vec{u}\|^2+\|\vec{v}\|^2-\|p_n\vec{u}-v\|^2\right]\\ &= \lim_{n\to\infty}\langle p_n\vec{u},\vec{v}\rangle\\ &= \lim_{n\to\infty}p_n\langle \vec{u},\vec{v}\rangle\\ &= \lambda\langle \vec{u},\vec{v}\rangle\,.\end{aligned}$$

Por ende,

$$\langle \lambda\vec{u},\vec{v}\rangle = \lambda\langle \vec{u},\vec{v}\rangle$$

para todo $\lambda\in\mathbb{R}$ y $\forall\vec{u},\vec{v}\in\mathbf{E}$.

De 1, 2, 3 y 4 $\langle\cdot,\cdot\rangle$ es un producto interior y la norma inducida por éste coincide con la norma del espacio, lo cual demuestra el teorema. ■

❍ **Nota 4.13** Observe que con el resultado precedente se prueba que una condición necesaria y también suficiente para que una norma en un espacio vectorial provenga de un producto interior, es que cumpla con la identidad del paralelogramo (4.41). En tal caso, el producto de donde proviene está definido por la identidad de polarización (4.42).

4.2.4 Normas equivalentes

El concepto de proximidad es relativo a la norma con la que se mide la distancia entre los vectores. No es lo mismo que dos funciones $f,g\in C[a,b]$ estén próximas con la norma $\|f-g\|=\left(\int_a^b (f(x)-g(x))^2\,dx\right)^{1/2}$ que con la norma uniforme $\|f(x)-g(x)\|_\infty$. Sin embargo, existen pares de normas para las que dos vectores son próximos respecto a una de ellas si y sólo si son próximos respecto a la otra norma. A este tipo de normas se les dice equivalentes y nos abocamos a su estudio en este apartado.

Definición 4.14 *Sean* $\mathbf{E}$ *un espacio vectorial y* $\|\cdot\|$, $\|\cdot\|'$ *un par de normas en* $\mathbf{E}$. *Diremos que la primera norma es equivalente a la segunda si existen* $\alpha,\beta\in\mathbb{R}$ *positivos tales que*

$$\alpha\,\|\vec{u}\|'\le\|\vec{u}\|\le\beta\,\|\vec{u}\|'\quad\forall\vec{u}\in\mathbf{E}.$$

En tal caso escribiremos $\|\vec{u}\|\sim\|\vec{u}\|'$.

[29] La continuidad de g en $\lambda\in\mathbb{R}$ equivale a $\lim_{n\to\infty}|f(x_n)-\lambda|=0$ para toda sucesión (x_n) tal que $\lim_{n\to\infty}|x_n-\lambda|=0$.

▶ **Ejemplo 4.54** En $\mathbb{R}^n$ la norma canónica $\|(x_1,x_2,\ldots,x_n)\| = \left(\sum_{k=1}^n x_k^2\right)^{1/2}$ es equivalente a la norma cúbica $\|(x_1,x_2,\ldots,x_n)\|_\infty = \text{máx}_{1\le k\le n}|x_k|$. En efecto, si $\vec{u} = (x_1,x_2,\ldots,x_n)$, entonces, puesto que $|x_k| \le \|\vec{u}\|_\infty\, \forall \vec{u} \in \mathbf{E}$, se tiene

$$\begin{aligned}\|\vec{u}\| &= \left(\sum_{k=1}^n x_k^2\right)^{1/2}\\ &\le \left(\sum_{k=1}^n \|\vec{u}\|_\infty^2\right)^{1/2}\\ &= \sqrt{n}\,\|\vec{u}\|_\infty.\end{aligned}$$

Por otra parte, dado que $\|\vec{u}\|_\infty^2 \le \sum_{k=1}^n |x_k|^2$,

$$\|\vec{u}\|_\infty \le \left(\sum_{k=1}^n \|\vec{u}\|_\infty^2\right)^{1/2}.$$

Por tanto,

$$\|\vec{u}\|_\infty \le \|\vec{u}\| \le \sqrt{n}\,\|\vec{u}\|_\infty \quad \forall \vec{u} \in \mathbf{E}.$$

Luego,

$$\|\vec{u}\| \sim \|\vec{u}\|_\infty. ◀$$

El siguiente teorema es sencillo de probar y su demostración se deja de ejercicio al lector.

Teorema 4.20 *Sea* $\mathbf{E}$ *un espacio vectorial y* $\|\cdot\|$, $\|\cdot\|'$, $\|\cdot\|''$ *normas en este espacio. Entonces:*

1. $\|\cdot\| \sim \|\cdot\|$ (Reflexividad).
2. $\|\cdot\| \sim \|\cdot\|' \Rightarrow \|\cdot\|' \sim \|\cdot\|$ (Simetría).
3. $\|\cdot\| \sim \|\cdot\|'$ *y* $\|\cdot\|' \sim \|\cdot\|'' \Rightarrow \|\cdot\| \sim \|\cdot\|''$ (Transitividad).

Así, del teorema anterior, si una norma es equivalente a una segunda, por simetría ésta será equivalente a la primera. Por esta razón, de aquí en adelante, diremos que dos normas son equivalentes si satisfacen la definición 4.14.

Sean ahora $\|\cdot\|$, $\|\cdot\|'$ un par de normas equivalentes en el espacio $\mathbf{E}$, $\vec{u}_0 \in \mathbf{E}$ y $B(\vec{u}_0,r_1)$ y $B'(\vec{u}_0,r_1)$ bolas abiertas respecto a sendas normas. Sean α, β números positivos tales

$$\alpha\,\|\vec{u}\|' \le \|\vec{u}\| \le \beta\,\|\vec{u}\|' \quad \forall \vec{u} \in \mathbf{E}.$$

1. Sea $r = r_1/\beta$, entonces

$$\begin{aligned}\vec{v} \in B'(\vec{u}_0,r) &\Rightarrow \|\vec{u}_0 - \vec{v}\|' < r\\ &\Rightarrow \|\vec{u}_0 - \vec{v}\| \le \beta\,\|\vec{u}_0 - \vec{v}\|' < \beta r = \beta\frac{r_1}{\beta} = r_1\\ &\Rightarrow \vec{v} \in B(\vec{u}_0,r_1)\\ &\Rightarrow B'(\vec{u}_0,r) \subset B(\vec{u}_0,r_1).\end{aligned}$$

2. Sea $r' = \alpha r_2$, entonces

$$\begin{aligned}\vec{v} \in B(\vec{u}_0, r') &\Rightarrow \|\vec{u}_0 - \vec{v}\| < r' \\ &\Rightarrow \|\vec{u}_0 - \vec{v}\|' \le \frac{1}{\alpha}\|\vec{u}_0 - \vec{v}\| < \frac{1}{\alpha} r' = \frac{1}{\alpha}\alpha r_2 = r_2 \\ &\Rightarrow \vec{v} \in B'(\vec{u}_0, r_2) \\ &\Rightarrow B(\vec{u}_0, r') \subset B'(\vec{u}_0, r_2).\end{aligned}$$

Con lo cual, hemos probado el siguiente teorema que justifica la discusión dada al inicio de esta subsección respecto a la proximidad entre vectores relativa a distintas normas.

Teorema 4.21 *Sea* **E** *un espacio vectorial y* $\|\cdot\|$, $\|\cdot\|'$ *un par de normas equivalentes en este espacio. Entonces:*

1. *Toda bola* $B(\vec{u}_0, r_1)$ *relativa a la norma* $\|\cdot\|$ *contiene una bola* $B'(\vec{u}_0, r)$ *relativa a la norma* $\|\cdot\|'$.
2. *Toda bola* $B'(\vec{u}_0, r_2)$ *relativa a la norma* $\|\cdot\|'$ *contiene una bola* $B(\vec{u}_0, r)$ *relativa a la norma* $\|\cdot\|$.

La figura 4-28 ilustra este teorema en el caso de las normas equivalentes del ejemplo 4.54 para $n = 2$.

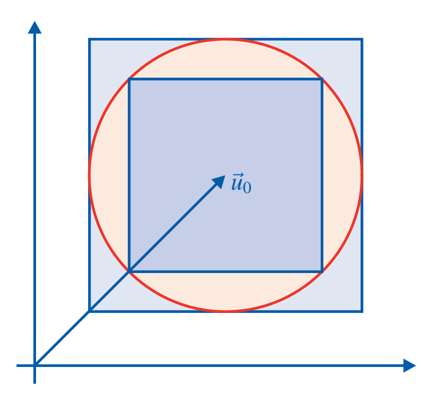


Figura 4-28 • Toda bola relativa a la norma canónica contiene una bola del mismo centro relativa a la norma cúbica y viceversa.

► **Ejemplo 4.55** Determinar si las normas

$$\|f\|_1 = \int_0^1 |f(x)|dx$$

y la norma $\|f\|_\infty = \text{máx}_{0 \le x \le 1} |f(x)|$ son equivalentes en $C[0,1]$.◄

Solución Supongamos que existen $\alpha > 0$ y $\beta > 0$, un par de números reales tales

$$\alpha \|f\|_\infty \le \|f\|_1 \le \beta \|f\|_\infty \quad \forall f \in C[0,1].$$

Sea, para cada número entero positivo m, la función continua $f_m(x) = e^{-mx}$, $0 \le x \le 1$. Entonces, para cada m,

$$\|f_m\|_\infty = \max_{0\le x\le 1} |e^{-mx}| = e^{-m\cdot 0} = 1 \quad \text{y}$$

$$\begin{aligned}\|f_m\|_1 &= \int_0^1 |f_m(x)|dx \\ &= \int_0^1 e^{-mx}dx \\ &= \frac{1}{m}\left(1-e^{-m}\right).\end{aligned}$$

Por tanto,

$$\alpha \cdot 1 \le \frac{1}{m}\left(1-e^{-m}\right) \quad \forall m = 1,2,\ldots$$

Pero, el lado derecho de la precedente desigualdad tiende a 0 cuando m tiende a infinito. Luego, existe m_0 tal que $\frac{1}{m_0}\left(1-e^{-m_0}\right) < \alpha$, tendríamos entonces

$$\alpha \le \frac{1}{m_0}\left(1-e^{-m_0}\right) < \alpha,$$

lo cual es una contradicción. Por tanto, estas normas no pueden ser equivalentes.[30] ✔

El ejemplo precedente muestra que no todas las normas son equivalentes. La dificultad intrínseca de este ejemplo radica en el hecho de que la dimensión del espacio $C[0,1]$ es infinta. Sin embargo, en espacios de dimensión finita, la situación cambia completamente; pues en ellos resulta ser que todas las normas son equivalentes. Este importantísimo resultado lo hacemos patente en el siguiente teorema, aunque no lo demostraremos aquí; e invitamos al lector a consultar, si así lo desea, la demostración dada en el apéndice C.

Teorema 4.22 ***(Equivalencia de normas en dimensión finita)*** *Si* **E** *un espacio vectorial, entonces cualquier par de normas en* **E** *son equivalentes; es decir, en un espacio de dimensión finita todas las normas en él son equivalentes.*

La proximidad de vectores es un concepto fundamental de las matemáticas. A través de la idea de proximidad es como se pueden definir límites, continuidad, derivación, diferenciación, derivadas parciales, optimización (máximos y mínimos locales) para funciones de una y de varias variables; convergencia de sucesiones y series; y conceptos geométricos avanzados que tienen que ver con una importante rama de las matemáticas llamada topología (particularmente de los espacios vectoriales normados en nuestro caso); etc. En todos estos temas de estudio, la proximidad depende de las normas con la que se trabaje; sin embargo, todo aquello que, en términos de proximidad, valga para ciertas normas, seguirá siendo válido si éstas se cambian por sendas normas equivalentes. Por ello, el teorema 4.22 bien

[30]Note que en este caso sí existe β tal que $\|f\|_1 \le \beta\|f\|_\infty$ $\forall f \in C[a,b]$; pues, dado que $|f(x)| \le \|f\|_\infty$ para todo $x \in [a,b]$, se tiene $\int_0^1 |f(x)|dx \le \int_0^1 \|f\|_\infty dx = (b-a)\|f\|_\infty$. Luego $\beta = b-a$ es un escalar que funciona para este fin.

podría ser llamado el *teorema fundamental de espacios vectoriales normados de dimensión finita*, pues significa que en estos espacios se puede trabajar con las normas que se deseen, evidentemente las más cómodas, pues todo lo que sea válido para ellas (en términos de proximidad) será valido para cualesquiera otras normas de estos espacios. Así que el lector debe tener muy en cuenta esto, utilizar las normas que más convengan y no necesariamente trabajar con normas que, por ser históricas en las matemáticas o por desconocimiento del teorema 4.22, se utilizan con gran frecuencia como es el caso de la norma euclidiana, la norma canónica en el espacio $\mathbb{R}^n$.

Normas p

Terminamos este apartado con un ejemplo de normas que son generalización directa de la norma canónica en $\mathbb{R}^n$, las normas $\|\cdot\|_p$ ($p \geq 1$). Este ejemplo bien podría haberse dado al inicio de esta subsección; pero por la dificultad que entraña se decidió ponerlo al final. Además, es una buena excusa para ilustrar el concepto de equivalencia de normas en $\mathbb{R}^n$ en este extenso conjunto de normas y explicar el porqué de la notación $\|\cdot\|_\infty$ para la norma cúbica.

Definición 4.15 *Sea $p \geq 1$ un número real y $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Se define*

$$\|\vec{u}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

A $\|\vec{u}\|_p$ se le llama norma p del vector $\vec{u}$.

❍ **Nota 4.14** Observe que si $p = 2$, se obtiene la norma canónica en $\mathbb{R}^n$. Y si $p = 1$ se obtiene la norma $\|\cdot\|_1$ que estudiamos en el ejemplo 4.42.

Antes de mostrar que $\|\vec{u}\|_p$ es efectivamente una norma, necesitamos probar algunos resultados preliminares. En el lema 3.1 (cfr. pág. 123) vimos que

$$2ab \leq a^2 + b^2$$

para cualquier par de números reales a y b. El lema 4.2 es una generalización de este resultado que utilizaremos para demostrar la desigualdad triangular de las normas p.

Definición 4.16 *Sean $p > 1$ y $p^* > 1$ números reales. Se dice que p y p^* son* ***índices conjugados*** *si*

$$\frac{1}{p} + \frac{1}{p^*} = 1.$$

► **Ejemplo 4.56**

- $p = 2$ y $p^* = 2$ son índices conjugados.
- $p = 3$ y $p^* = 3/2$ son índices conjugados.◄

Las siguientes propiedades de índices conjugados, contenidas en el lema 4.1, son fáciles de probar y su demostración se deja de ejercicio al lector.

Lema 4.1 *Sean p y p^* índices conjugados. Entonces:*

1. $\dfrac{1}{p^*} = 1 - \dfrac{1}{p}$.
2. $\dfrac{1}{p} = 1 - \dfrac{1}{p^*}$.
3. $p^*(p-1) = p$.
4. $p(p^*-1) = p^*$.

Lema 4.2 *Sean $a, b \in \mathbb{R}$ un par de números positivos y p, p^* índices conjugados. Entonces,*

$$a^{1/p} b^{1/p^*} \leq \frac{a}{p} + \frac{b}{p^*} \tag{4.47}$$

DEMOSTRACIÓN ■ Sea $\alpha \in \mathbb{R}$, $0 < \alpha < 1$, y sea $f : [0, \infty) \to \mathbb{R}$ la función definida por

$$f(x) = x^\alpha - \alpha x + \alpha.$$

Entonces,

$$f'(x) = \alpha(x^{\alpha-1} - 1) \quad \text{y}$$
$$f''(x) = \alpha(\alpha - 1)x^{\alpha-2} < 0 \text{ en } (0, \infty).$$

Por tanto, el máximo de f en $[0, \infty)$ se alcanza en $x = 1$ y, por ende,

$$x^\alpha - \alpha x + \alpha \leq f(1) = 1 \quad \forall x \in [0, \infty) \tag{4.48}$$

Sean $\alpha = 1/p$ y $x = a/b$. Entonces $0 < \alpha < 1$ y $a/b \in (0, \infty)$, por tanto, al sustituir estos valores en (4.48) se obtiene:

$$\frac{a^{1/p}}{b^{1/p}} - \frac{1}{p}\frac{a}{b} + \frac{1}{p} \leq 1 \tag{4.49}$$

Multipliquemos ambos lados de la desigualdad (4.49) por b para tener:

$$\frac{a^{1/p} b}{b^{1/p}} - \frac{a}{p} + \frac{b}{p} \leq b,$$

que equivale a

$$\begin{aligned} a^{1/p} b^{1-1/p} &\leq \frac{a}{p} + b - \frac{b}{p} \\ &= \frac{a}{p} + b\left(1 - \frac{1}{p}\right). \end{aligned}$$

Pero $1 - \dfrac{1}{p} = p^*$ (por el lema 4.1); por tanto, la desigualdad precedente se transforma en

$$a^{1/p} b^{1/p^*} \leq \frac{a}{p} + \frac{b}{p^*}$$

que es lo que se quería demostrar. ■

❍ **Nota 4.15** Si sustituimos el caso particular $p = 2$ en (4.47) se obtiene

$$a^{1/2}b^{1/2} \leq \frac{a}{2} + \frac{b}{2}$$

y al elevar al cuadrado ambos lados de la desigualdad anterior

$$ab \leq \frac{(a+b)^2}{4},$$

que equivale a

$$4ab \leq a^2 + 2ab + b^2;$$

esto es,

$$2ab \leq a^2 + b^2$$

Es decir, que la desigualdad (4.47) tiene como caso particular la desigualdad (3.5) del lema 3.1 (cfr. pág. 123); que se utilizó para probar la desigualdad de Schwarz (teorema 3.1, cfr. pág. 124).

El siguiente lema establece una generalización de la desigualdad de Schwarz en $\mathbb{R}^n$, la desigualdad de Hölder, que como en el caso de la desigualdad de Schwarz, servirá para demostrar la desigualdad triangular de la norma p.

Lema 4.3 *Sean $\vec{u} = (x_1, x_2, \ldots, x_n)$ y $\vec{v} = (y_1, y_2, \ldots, y_n)$ un par de vectores en $\mathbb{R}^n$ y p, p^* índices conjugados. Entonces*

1. $$\left|\sum_{i=1}^{n} x_i y_i\right| \leq \|\vec{u}\|_p \|\vec{v}\|_{p^*} \quad \textbf{(Desigualdad de Hölder)} \tag{4.50}$$

2. $$\|\vec{u}+\vec{v}\|_p \leq \|\vec{u}\|_p + \|\vec{v}\|_p \quad \textbf{(Desigualdad de Minkowski)} \tag{4.51}$$

DEMOSTRACIÓN ■ 1. Sea $i \in \{1, 2, \ldots, n\}$ un índice fijo por el momento y pongamos[31]

$$a = \frac{|x_i|^p}{\|\vec{u}\|_p^p} \qquad \text{y} \qquad b = \frac{|y_i|^{p^*}}{\|\vec{v}\|_{p^*}^{p^*}}$$

en la desigualdad (4.47) del lema 4.2, entonces

$$\frac{|x_i|}{\|\vec{u}\|_p} \frac{|y_i|}{\|\vec{v}\|_{p^*}} \leq \frac{1}{p} \frac{|x_i|^p}{\|\vec{u}\|_p^p} + \frac{1}{p^*} \frac{|y_i|^{p^*}}{\|\vec{v}\|_{p^*}^{p^*}} \quad \forall i = 1, 2, \ldots, n.$$

Al sumar todos los índices desde $i = 1$ hasta $i = n$ en la precedente desigualdad obtenemos

$$\sum_{i=1}^{n} \frac{1}{\|\vec{u}\|_p \|\vec{v}\|_{p^*}} |x_i||y_i| \leq \sum_{i=1}^{n} \left(\frac{1}{p} \frac{|x_i|^p}{\|\vec{u}\|_p^p} + \frac{1}{p^*} \frac{|y_i|^{p^*}}{\|\vec{v}\|_{p^*}^{p^*}} \right)$$

[31] Observe la analogía que hay con la demostración de la desigualdad de Schwarz (teorema 3.1, cfr. pág. 124).

que equivale a

$$\begin{aligned}\frac{1}{\|\vec{u}\|_p \|\vec{v}\|_{p^*}} \sum_{i=1}^n |x_i||y_i| &\le \frac{1}{p\|\vec{u}\|_p^p}\sum_{i=1}^n |x_i|^p + \frac{1}{p^*\|\vec{v}\|_{p^*}^{p^*}}\sum_{i=1}^n |y_i|^{p^*} \\ &= \frac{1}{p\|\vec{u}\|_p^p}\|\vec{u}\|_p^p + \frac{1}{p^*\|\vec{v}\|_{p^*}^{p^*}}\|\vec{v}\|_{p^*}^{p^*} \\ &= \frac{1}{p} + \frac{1}{p^*} = 1\end{aligned}$$

y, por tanto,

$$\sum_{i=1}^n |x_i||y_i| \le \|\vec{u}\|_p \|\vec{v}\|_{p^*} \tag{4.52}$$

Pero, puesto que $|x_i||y_i| = |x_iy_i|$ y por la desigualdad triangular para el valor absoluto de números reales, $|\sum_{i=1}^n x_iy_i| \le \sum_{i=1}^n |x_iy_i|$, la desigualdad (4.52) implica

$$\left|\sum_{i=1}^n x_iy_i\right| \le \|\vec{u}\|_p \|\vec{v}\|_{p^*}.$$

2. Para cada $i = 1, 2, \ldots, n$ se tiene

$$\begin{aligned}|x_i + y_i|^p &= |x_i + y_i||x_i + y_i|^{p-1} \\ &\le (|x_i| + |y_i|)\, |x_i + y_i|^{p-1} \\ &= |x_i||x_i + y_i|^{p-1} + |y_i||x_i + y_i|^{p-1}\end{aligned}$$

y, por tanto,

$$\sum_{i=1}^n |x_i + y_i|^p \le \sum_{i=1}^n |x_i||x_i + y_i|^{p-1} + \sum_{i=1}^n |y_i||x_i + y_i|^{p-1}.$$

Por (4.52), si $\vec{w} = \left(|x_1 + y_1|^{p-1}, |x_2 + y_2|^{p-1}, \ldots, |x_n + y_n|^{p-1}\right)$,

$$\sum_{i=1}^n |x_i||x_i + y_i|^{p-1} \le \|\vec{u}\|_p \|\vec{w}\|_{p^*} \quad \text{y}$$

$$\sum_{i=1}^n |y_i||x_i + y_i|^{p-1} \le \|\vec{v}\|_p \|\vec{w}\|_{p^*}$$

Por lo que

$$\begin{aligned}\sum_{i=1}^n |x_i + y_i|^p &\le \|\vec{u}\|_p \|\vec{w}\|_{p^*} + \|\vec{v}\|_p \|\vec{w}\|_{p^*} \\ &= \left(\|\vec{u}\|_p + \|\vec{v}\|_p\right)\|\vec{w}\|_{p^*}.\end{aligned}$$

Pero,

$$\begin{aligned}\|\vec{w}\|_{p^*} &= \left(\sum_{i=1}^n |x_i + y_i|^{(p-1)p^*}\right)^{1/p^*} \\ &= \left(\sum_{i=1}^n |x_i + y_i|^p\right)^{1/p^*}.\end{aligned}$$

Por tanto,

$$\sum_{i=1}^{n}|x_i+y_i|^p \le \left(\|\vec{u}\|_p+\|\vec{v}\|_p\right)\left(\sum_{i=1}^{n}|x_i+y_i|^p\right)^{1/p^*},$$

de donde

$$\frac{\sum_{i=1}^{n}|x_i+y_i|^p}{\left(\sum_{i=1}^{n}|x_i+y_i|^p\right)^{1/p^*}} \le \|\vec{u}\|_p+\|\vec{v}\|_p;$$

esto es:

$$\left(\sum_{i=1}^{n}|x_i+y_i|^p\right)^{1-1/p^*} \le \|\vec{u}\|_p+\|\vec{v}\|_p$$

que equivale (pues $1-1/p^*=1/p$) a

$$\left(\sum_{i=1}^{n}|x_i+y_i|^p\right)^{1/p} \le \|\vec{u}\|_p+\|\vec{v}\|_p.$$

Es decir,

$$\|\vec{u}+\vec{v}\|_p \le \|\vec{u}\|_p+\|\vec{v}\|_p. \quad ■$$

❍ **Nota 4.16** Observe que la desigualdad de Hölder (4.50) tiene como caso particular la desigualdad de Schwarz en $\mathbb{R}^n$; pues si sustituimos $p=2$ en ésta obtenemos $|\vec{u}\cdot\vec{v}| \le \|\vec{u}\|_2\,\|\vec{v}\|_2$; pero $\|\cdot\|_2$ es precisamente la norma canónica inducida por el producto punto de vectores en $\mathbb{R}^n$.

Estamos ya capacitados para demostrar que las normas p son efectivamente normas en $\mathbb{R}^n$, lo cual hacemos en el siguiente teorema.

Teorema 4.23 *Si $p \ge 1$ es un número real, entonces*

$$\|\vec{u}\|_p = \left(\sum_{n=1}^{n}|x_i|^p\right)^{1/p},$$

$\vec{u}=(x_1,x_2,\ldots,x_n)$, es una norma en $\mathbb{R}^n$.

DEMOSTRACIÓN ■ Si $p=1$, ya probamos esta afirmación en el ejemplo 4.42 (cfr. pág. 305). Supongamos entonces que $p>1$.

1. $\|\vec{u}\|_p = \left(\sum_{n=1}^{n}|x_i|^p\right)^{1/p} \ge 0 \quad \forall \vec{u}=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$.
2. Claramente $\|\vec{0}_{\mathbb{R}^n}\|_p=0$. Supongamos que $\|\vec{u}\|_p=0$, entonces

$$|x_i|^p \le \sum_{n=1}^{n}|x_i|^p = \|\vec{u}\|_p^p = 0 \Rightarrow x_i=0\ \forall i \Rightarrow \vec{u}=\vec{0}_{\mathbb{R}^n}.$$

3. Si $\lambda \in \mathbb{R}$ y $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, entonces

$$\begin{aligned} \|\vec{u}\|_p &= \left(\sum_{n=1}^{n} |\lambda x_i|^p\right)^{1/p} \\ &= \left(|\lambda|^p \sum_{n=1}^{n} |x_i|^p\right)^{1/p} \\ &= |\lambda| \left(\sum_{n=1}^{n} |x_i|^p\right)^{1/p} \\ &= |\lambda| \, \|\vec{u}\|_p \,. \end{aligned}$$

4. La desigualdad triangular ya se probó en el lema 4.3 (desigualdad de Minkowski (4.51)). ■

▶ **Ejemplo 4.57** Mostrar, sin utilizar el teorema de equivalencia de normas en espacios de dimensión finita, que si $p > 1$, entonces

$$\|\cdot\|_p \sim \|\cdot\|_\infty$$

en el espacio euclidiano $\mathbb{R}^n$.◀

DEMOSTRACIÓN ■ Sea $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, entonces

$$|x_i|^p \le \|\vec{u}\|_\infty^p \quad \forall i = 1, 2, \ldots, n.$$

Por tanto,

$$\sum_{i=1}^{n} |x_i|^p \le \sum_{i=1}^{n} \|\vec{u}\|_\infty^p = n \, \|\vec{u}\|_\infty^p \,.$$

Luego,

$$\left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \le n^{1/p} \, \|\vec{u}\|_\infty \tag{4.53}$$

Por otra parte, es claro que

$$\|\vec{u}\|_\infty^p \le \sum_{i=1}^{n} |x_i|^p,$$

de donde

$$\|\vec{u}\|_\infty \le \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \tag{4.54}$$

De (4.53) y (4.54) se desprende

$$\|\vec{u}\|_\infty \le \|\vec{u}\|_p \le n^{1/p} \, \|\vec{u}\|_\infty \tag{4.55}$$

Y, por ende, $\|\cdot\|_p \sim \|\cdot\|_\infty$ con $\alpha = 1$ y $\beta = n^{1/p}$. ■

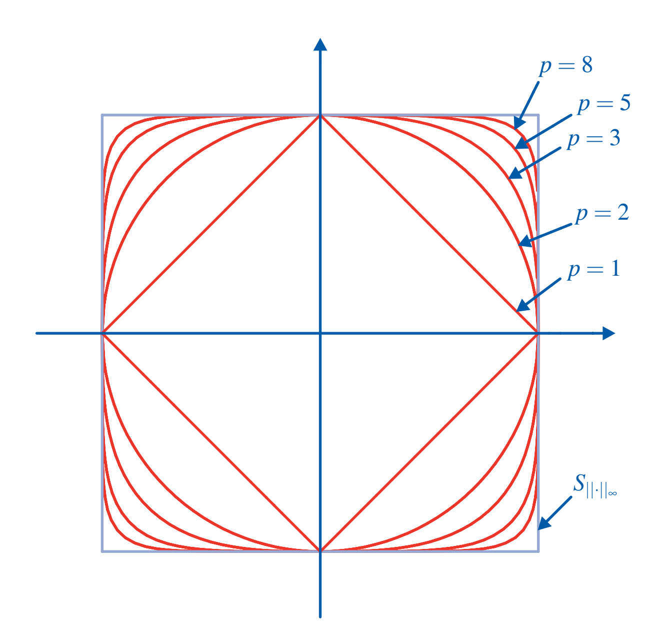


Figura 4-29 • La norma p de cualquier vector tiende a la norma cúbica de éste cuando p toma valores cada vez más grandes. En esta figura hemos graficado las esferas $S(\vec{0},r)$ relativas a la norma $p=1,2,3,5,8$ y la esfera $S_{\|\cdot\|_\infty}$ de mismo centro y radio relativa a la norma cúbica. Observe cómo las esferas para las normas p son cada vez más cercanas a la esfera para la norma cúbica en la medida en que p crece.

Ahora estamos listos para explicar la razón de la notación $\|\cdot\|_\infty$ para la norma cúbica. Para ello, fijemos un vector arbitrario $\vec{u}\in\mathbb{R}^n$ y calculemos

$$\lim_{p\to\infty}\|\vec{u}\|_p\,.$$

Por (4.55) se tiene

$$\|\vec{u}\|_\infty \le \lim_{p\to\infty}\|\vec{u}\|_p \le \lim_{n\to\infty} n^{1/p}\,\|\vec{u}\|_\infty = 1\cdot\|\vec{u}\|_\infty\,;$$

de donde

$$\lim_{p\to\infty}\|\vec{u}\|_p = \|\vec{u}\|_\infty\,.$$

Esto es, la norma p tiende a la norma cúbica cuando p tiende a infinito. Debido a este hecho es que tradicionalmente a la norma cúbica se le denota con el símbolo $\|\cdot\|_\infty$. En la figura 4-29, hemos graficado las esferas de centro $(0,0)$ relativas a la norma p para algunos valores crecientes de p y la esfera de mismo centro relativa a la norma cúbica, todas con el mismo radio r. En ella se puede observar cómo las p-esferas se aproximan al cuadrado de arista $2r$ con centro en el origen (la esfera para la norma cúbica) conforme p aumenta.

4.2.5 Construcción de normas en espacios de dimensión finita a partir de normas en $\mathbb{R}^n$

En esta breve subsección veremos cómo es posible, de manera natural, construir normas en un espacio de dimensión finita a través de normas en $\mathbb{R}^n$. Esta construción siempre dependerá de la base con la

que se estén describiendo los vectores del espacio. En realidad, el objetivo de este apartado es ahorrar tiempo al lector para que no tenga que demostrar que las normas que se definen por analogía directa de normas en $\mathbb{R}^n$ a espacios de dimensión finita, son efectivamente normas en estos espacios. Sean $\mathbf{E}$ un espacio vectorial de dimensión finita y $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ una base de este espacio; entonces si $\vec{u} \in \mathbf{E}$, existen escalares $a_1, a_2, \ldots, a_n$ tales que

$$\vec{u} = a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n \tag{4.56}$$

Si $b_1, b_2, \ldots, b_n$ son escalares tales que también

$$\vec{u} = b_1\vec{e}_1 + b_2\vec{e}_2 + \cdots + b_n\vec{e}_n,$$

entonces,

$$(a_1 - b_1)\vec{e}_1 + (a_2 - b_2)\vec{e}_1 + \cdots + (a_n - b_n)\vec{e}_1 = \vec{0}_{\mathbf{E}}$$

y puesto que los vectores $\vec{e}_i$ son L.I., se desprende que

$$a_i = b_i \quad \forall i = 1, 2, \ldots, n.$$

Luego, todo vector $\vec{u} \in \mathbf{E}$ se puede escribir como combinación lineal de los vectores $\vec{e}_i$ como en (4.56) de manera única. Si convenimos en que el orden en los términos de (4.56) es el mismo para cualquier vector (el primer término es un escalar que multiplica al primer vector $\vec{e}_1$, el segundo término un escalar por el vector $\vec{e}_2$, etc., y el último término es un escalar que multiplica al vector $\vec{e}_n$), diremos que la base está ordenada y escribiremos $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ en lugar de $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ para subrayar este hecho.[32] En tal caso a $(a_1, a_2, \ldots, a_n)$ le llamaremos el **vector de coordenadas** del vector $\vec{u}$ **relativo** a la base ordenada $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$.

► **Ejemplo 4.58** Sea la base ordenada $(3, x-1, 2x^2)$ del espacio de polinomios $\mathbf{P}_2$. Encontrar el vector de coordenadas del polinomio $p(x) = 1 - 3x + 4x^2$ relativo a esta base ordenada.◄

Solución Busquemos escalares a_1, a_2, a_3, tales que

$$a_1 \cdot 3 + a_2 \cdot (x-1) + a_3(2x^2) = 1 - 3x + 4x^2;$$

entonces,

$$3a_1 - a_2 + a_2x + 2a_3x^2 = 1 - 3x + 4x^2$$

y, por tanto,

$$\begin{aligned} 3a_1 - a_2 &= 1, \\ a_2 &= -3, \\ 2a_3 &= 4; \end{aligned}$$

[32]Cfr. la discusión que precede al teorema 4.9 en la página 265.

de donde

$$a_2 = -3, a_3 = 2, a_1 = -\frac{2}{3}.$$

Por lo que el vector de coordenadas del polinomio $p(x)$ relativo a esta base ordenada es

$$(-2/3, -3, 2). \quad ✔$$

Sea ahora $\|\cdot\|$ una norma en $\mathbb{R}^n$ (no necesariamente la canónica) y $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ una base ordenada del espacio **E**. Entonces, si $\vec{u} \in \mathbf{E}$ y $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ es el vector de coordenadas de $\vec{u}$ relativo a la base ordenada, definimos

$$\|\vec{u}\|_{\mathbf{E}} = \|(a_1, a_2, \ldots, a_n)\|.$$

Afirmamos que ésta es una norma en **E**. En efecto:

1. $\|\vec{u}\|_{\mathbf{E}} = \|(a_1, a_2, \ldots, a_n)\| \geq 0 \quad \forall \vec{u} \in \mathbf{E}$.
2. $\vec{0}_{\mathbf{E}} = 0 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + \cdots + 0 \cdot \vec{e}_n$, por tanto, $\|\vec{0}_{\mathbf{E}}\|_{\mathbf{E}} = \|\vec{0}_{\mathbb{R}^n}\| = 0$. $\|\vec{u}\|_{\mathbf{E}} = 0 \Rightarrow \|(a_1, a_2, \ldots, a_n)\| = 0 \Rightarrow (a_1, a_2, \ldots, a_n) = \vec{0}_{\mathbb{R}^n} \Rightarrow \vec{u} = a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n = \vec{0}_{\mathbf{E}}$.
3. Si $\lambda \in \mathbb{R}$ y $\vec{u} = a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n \in \mathbf{E}$, entonces

$$\begin{aligned}
\|\lambda\vec{u}\|_{\mathbf{E}} &= \|\lambda(a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n)\|_{\mathbf{E}} \\
&= \|(\lambda a_1)\vec{e}_1 + (\lambda a_2)\vec{e}_2 + \cdots + (\lambda a_n)\vec{e}_n\|_{\mathbf{E}} \\
&= \|\lambda(a_1, a_2, \ldots, a_n)\| \\
&= |\lambda| \, \|(a_1, a_2, \ldots, a_n)\| \\
&= |\lambda| \, \|\vec{u}\|_{\mathbf{E}}.
\end{aligned}$$

4. Si $\vec{u}, \vec{v} \in \mathbf{E}$ y $(a_1, a_2, \ldots, a_n)$, $(b_1, b_2, \ldots, b_n)$ son los vectores de coordenadas de $\vec{u}$ y $\vec{v}$, respectivamente, entonces

$$\begin{aligned}
\|\vec{u} + \vec{v}\|_{\mathbf{E}} &= \|(a_1 + b_1)\vec{e}_1 + (a_2 + b_2)\vec{e}_2 + \cdots + (a_n + b_n)\vec{e}_n\|_{\mathbf{E}} \\
&= \|(a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)\| \\
&= \|(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n)\| \\
&\leq \|(a_1, a_2, \ldots, a_n)\| + \|(b_1, b_2, \ldots, b_n)\| \\
&= \|\vec{u}\|_{\mathbf{E}} + \|\vec{v}\|_{\mathbf{E}}.
\end{aligned}$$

▶ **Ejemplo 4.59** Sea el caso dado en el ejemplo 4.58 y la norma canónica en $\mathbb{R}^n$, $\|\cdot\|$. Entonces

$$\begin{aligned}
\|p(x)\|_{\mathbf{P}_2} &= \|(-2/3, -3, 2)\| \\
&= \frac{11}{3}.
\end{aligned}$$

Si se toma $\|\cdot\|_\infty$ en lugar de la canónica se tiene

$$\begin{aligned}
\|p(x)\|_{\mathbf{P}_2} &= \|(-2/3, -3, 2)\|_\infty \\
&= 3. ◀
\end{aligned}$$

Recíprocamente, toda norma $\|\cdot\|_{\mathbf{E}}$ del espacio $\mathbf{E}$ produce una norma en $\mathbb{R}^n$; pues si $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ y $\vec{u} = a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n$, entonces la aplicación

$$(a_1, a_2, \ldots, a_n) \mapsto \|(a_1, a_2, \ldots, a_n)\| = \|\vec{u}\|_{\mathbf{E}}$$

en una norma en $\mathbb{R}^n$.

4.2.6 Aproximaciones óptimas en espacios normados

Aquí veremos cómo es posible encontrar aproximaciones óptimas en espacios normados como lo hicimos en espacios con producto interior; aunque dicha norma no necesariamente provenga de un producto escalar.

Definición 4.17 ***(Aproximaciones óptimas)*** *Sean* $\mathbf{E}$ *un espacio vectorial normado con norma* $\|\cdot\|$, $\vec{u}$ *un vector dado de* $\mathbf{E}$ *y* S *un subespacio. Decimos que* $\vec{p}^*$ *es* ***aproximación óptima*** *de* $\vec{u}$ *en* S *si*

$$\left\|\vec{u} - \vec{p}^*\right\| \leq \|\vec{u} - \vec{v}\| \quad \forall \vec{v} \in S$$

o, de manera equivalente, si

$$\min_{\vec{v} \in S} \|\vec{u} - \vec{v}\| = \left\|\vec{u} - \vec{p}^*\right\|.$$

El siguiente teorema garantiza la existencia de aproximaciones óptimas en el caso de ser S un subespacio de dimensión finita. Su demostración requiere de algunos conceptos básicos de funciones continuas en espacios vectoriales normados y la postergaremos al apéndice C. Invitamos al lector a que la consulte en el momento que desee y recomendamos su lectura, pues tiene aspectos muy interesantes y puede ser de gran provecho el intentar comprenderlos, al menos *grosso modo*.

Teorema 4.24 ***(Aproximaciones óptimas en espacios normados)*** *Sean* $\mathbf{E}$ *un espacio vectorial normado con norma* $\|\cdot\|$, S *un subespacio de dimensión finita en* $\mathbf{E}$ *y* $\vec{u}$ *un vector dado de* $\mathbf{E}$. *Entonces, existe una aproximación óptima* $\vec{p}^*$ *de* $\vec{u}$ *en* S.

Aproximaciones óptimas en $C[a,b]$ con la norma uniforme

Sean $f \in C[a,b]$ y $S = \mathbf{P}_n = \text{gn}(1, x, x^2, \ldots, x^n)$; por el teorema 4.24 existe un polinomo $p_n^*(x)$ de grado a lo más n tal que

$$\|f - p_n^*\|_\infty \leq \|f - p_n\|_\infty$$

para todo polinomio p_n de grado a lo más n. Esto ocurre para cada $n \in \mathbb{N}$; como es natural, surge la cuestión de si la sucesión de aproximaciones óptimas p_n^*, $n = 1, 2, \ldots$, converge a f respecto a la norma uniforme; esto es,

$$\lim_{n\to\infty} \|f - p_n^*\|_\infty = 0.$$

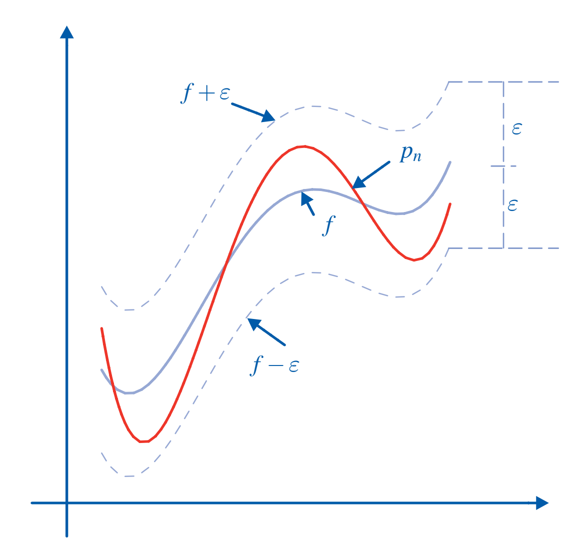


Figura 4-30 • Si $f \in C[a,b]$, para cualquier bola abierta $B(f,\varepsilon)$, relativa a la norma uniforme $\|\cdot\|_\infty$, existe un polinomio de grado n dentro de ella.

Recordemos que en el caso de las aproximaciones óptimas de polinomios trigonométricos a funciones continuas vimos que esta convergencia se daba en promedio cuadrático; esto es, para la norma $\|f-g\| = \left(\int_a^b (f-g)^2\right)^{1/2}$. Para las aproximaciones óptimas, con la norma uniforme, esto también sucede. El teorema que garantiza este hecho es uno de los más célebres e importantes resultados en las matemáticas. A continuación lo enunciamos y haremos su demostración plausible en la discusión posterior.

Teorema 4.25 ***(De aproximación de Weierstrass)*** *Sea $f \in C[a,b]$. Entonces, para cada $\varepsilon > 0$ existe un polinomio p_n de grado n (que depende de ε) tal que*

$$\|f-p_n\|_\infty < \varepsilon \tag{4.57}$$

La interpretación geométrica de este teorema viene ilustrada en la figura 4-30. En ella se muestra una función continua f en un intervalo $[a,b]$, una bola de centro f y radio ε respecto a la norma uniforme $\|\cdot\|_\infty$ y un polinomio p_n dentro de esta bola; por lo cual (4.57) se cumple para este polinomio.

❍ Nota 4.17

1. El teorema de aproximación de Weierstrass equivale a que, para cada función $f \in C[a,b]$, existe una sucesión de polinomios (p_n) que converge uniformemente a f, i.e.,

$$\lim_{n\to\infty} \|f-p_n\|_\infty = 0 \tag{4.58}$$

En efecto, si (p_n) es una sucesión de polinomios tal que $\lim_{n\to\infty}\|f-p_n\|_\infty = 0$ y $\varepsilon > 0$ es dado, entonces, por definición de límite, existe $n_0 \in \mathbb{N}$ tal que

$$\left\|f-p_{n_0}\right\|_\infty < \varepsilon$$

y se cumple entonces (4.57) del teorema de Weierstrass. Supongamos inversamente que se cumple (4.57) del teorema 4.25, entonces para cada $n \in \mathbb{N}$, existe un polinomio p_n tal que

$$\|f - p_n\|_\infty < \frac{1}{n}$$

de donde se desprende que (4.58) es válida.

2. Si $\varepsilon > 0$ es dado, por el teorema aproximación de Weierstrass existe un polinomio de grado $n_0 \in \mathbb{N}$ tal que

$$\left\|f - p_{n_0}\right\|_\infty < \varepsilon.$$

Sea p_n^* una sucesión de aproximaciones óptimas para la función f (cuya existencia está garantizada por el teorema 4.24), entonces

$$\left\|f - p_{n_0}^*\right\|_\infty \leq \left\|f - p_{n_0}\right\|_\infty < \varepsilon.$$

Dado que $\mathbf{P}_n \subset \mathbf{P}_m$ si $n \leq m$, se desprende que

$$\|f - p_m^*\| \leq \|f - p_n^*\|.$$

Por lo que,

$$\|f - p_n^*\|_\infty \leq \left\|f - p_{n_0}^*\right\|_\infty < \varepsilon \quad \forall n \geq n_0.$$

Por tanto,

$$\lim_{n\to\infty} \|f - p_n^*\|_\infty = 0$$

como afirmamos en la discusión ulterior al teorema 4.24.

Aproximación por polinomios de Bernstein

Es posible hacer una demostración constructiva del teorema de aproximación de Weierstrass haciendo explícita la sucesión de polinomios que converge uniformemente a la función dada. Esto se realiza mediante los llamados polinomios de Bernstein.

Definición 4.18 *Sean $f \in C[0,1]$ y $n \in \mathbb{N}$. Se define el n-ésimo polinomio de Bernstein para la función f como*

$$B_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}$$

donde

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

es el usual coeficiente binomial.

▶ **Ejemplo 4.60** Sea la función continua en $[0,1]$,

$$f(x) = x\cos(20x) + 4;$$

entonces:

1.
$$\begin{aligned}B_1(x) &= \sum_{k=0}^{1} f\left(\frac{k}{1}\right)\binom{1}{k}x^k(1-x)^{1-k}\\ &= f(0)\binom{1}{0}x^0(1-x)^1 + f(1)\binom{1}{1}x^1(1-x)^0\\ &= 4(1-x)+(\cos(20)+4)x\\ &= 4+\cos(20)x.\end{aligned}$$

2.
$$\begin{aligned}B_2(x) &= \sum_{k=0}^{2} f\left(\frac{k}{2}\right)\binom{2}{k}x^k(1-x)^{2-k}\\ &= f(0)\binom{2}{0}x^0(1-x)^2 + f\left(\frac{1}{2}\right)\binom{2}{1}x^1(1-x)^1\\ &\quad +f(1)\binom{2}{2}x^2(1-x)^0\\ &= 4(1-x)^2+2\left(\frac{1}{2}\cos(10)+4\right)x(1-x)\\ &\quad +(\cos(20)+4)x^2\\ &= 4-(\cos 10)x+(\cos 20-\cos 10)x^2.\blacktriangleleft\end{aligned}$$

En la figura 4-31 se muestran aproximaciones sucesivas para la función $f(x)=x\cos(20x)+4$ mediante polinomios de Bernstein B_n para los valores $n=30$, 50, 100, 150, 180, 250 y $n=260$. En ella se puede observar cómo $\|f-B_n\|_\infty \to 0$, debido a que el máximo de $|f(x)-B_n(x)|$ va tendiendo a cero en la medida que n aumenta.

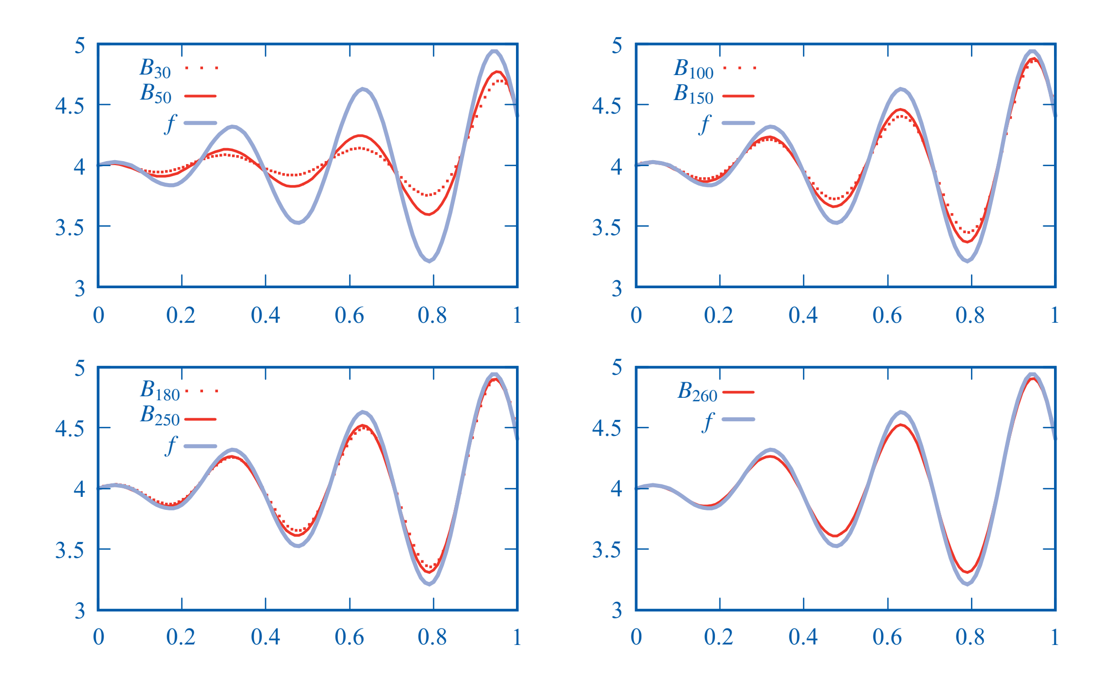


Figura 4-31 • Aproximaciones sucesivas de polinomios de Bernstein B_n, para diversos valores n (30, 50, 100, 150, 180, 250 y 260), a la función $f(x)=x\cos(20x)+4$.

❍ **Nota 4.18** Los polinomios de Bernstein aproximan uniformemente a funciones continuas en el intervalo $[0,1]$ y el teorema de Weierstrass afirma que este tipo de aproximación uniforme se puede efectuar en cualquier intervalo $[a,b]$. Pero en realidad, bajo una simple translación, se pueden construir los polinomios de Bernstein para funciones continuas en un intervalo $[a,b]$. En efecto, si $f \in C[a,b]$, sea $x = g(t) = (b-a)t + a$, $0 \le t \le 1$, entonces la función $f(x) = f(g(t))$ es una función continua en $[0,1]$ y los polinomios de Bernstein $B_n(x) = B_n(g(t))$ aproximan uniformemente a $f(g(t))$ en $[0,1]$ y, por tanto, aproximan uniformemente a $f(x)$ en $[a,b]$.

4.2.7 ¿Qué norma utilizar?

Cuando se busca la aproximación óptima de un elemento de un espacio normado mediante elementos de un subespacio, dicha aproximación depende de la norma con la que se esté trabajando. La elección de la norma en el espacio depende a su vez del tipo de problema que se desee resolver. Con el propósito de aclarar esta situación, planteamos a continuación dos ejemplos clásicos, en donde veremos cómo se elige la norma en una situación específica, y, de paso, por qué se definen de manera natural, en problemas de aplicación, dos de las normas más importantes en el espacio $C[a,b]$.

Problema de Chebichev

P. Lvovich Chebichev fue uno de los más grandes matemáticos rusos del siglo XIX; aparte de haber sido el creador de las bases de varias disciplinas matemáticas, las cuales aún están en pleno desarrollo, fue un notable ingeniero. Uno de sus intereses como tal, fue la construcción de un mecanismo capaz de reproducir el movimiento de una trayectoria dada. Sea $y = f(x)$, $a \le x \le b$, la trayectoria que se pretende reproducir. Se desea construir, bajo ciertos requisitos técnicos, un mecanismo tal que uno de sus dispositivos describa esta curva tan exactamente como sea posible cuando entre en funcionamiento. Chebichev, como buen ingeniero, construyó inicialmente un mecanismo que obtuviera una aproximación burda de la trayectoria dada. Así, un dispositivo de este mecanismo inicial describirá una curva

$$y = \varphi(x) \tag{4.59}$$

parecida a la curva dada $y = f(x)$ como se ilustra en la figura 4-32. Este primer mecanismo consta de ciertos dispositivos (engranes, palancas, etc.) todos ellos con medidas específicas, α_1, $\alpha_2, \ldots, \alpha_n$, que lo describen completamente y, por tanto, a la curva (4.59); son estos números los parámetros del mecanismo y de esta curva. Mejorar la aproximación del mecanismo significa modificar los parámetros α_i para ese fin. Aquí fue donde Chebichev abordó el problema matemáticamente. Para ello, consideró los posibles mecanismos que describirían curvas

$$y = \varphi(x, \alpha_1, \alpha_2, \ldots, \alpha_n) \tag{4.60}$$

para cada conjunto de parámetros α_i, $i = 1, 2, \ldots, n$. Cada una de éstas, representa una aproximación de la trayectoria $y = f(x)$. Sea $y = \varphi(x)$ la curva descrita por (4.60), donde por simplicidad hemos omitido los parámetros α_i, y sea

$$m = m(\alpha_1, \alpha_2, \ldots, \alpha_n) = \max_{a \le x \le b} |f(x) - \varphi(x)|$$

Figura 4-32 • Una trayectoria dada $y = f(x)$ y una burda aproximación de la misma, $y = \varphi(x)$; esta última está perfectamente determinada por los parámetros α_i del mecanismo que la produce.

como se ilustra en la figura 4-32. Entonces, dado que

$$|f(x) - \varphi(x)| \leq \max_{a \leq x \leq b} |f(x) - \varphi(x)| = m(\alpha_1, \alpha_2, \ldots, \alpha_n)$$

se tiene

$$|f(x) - \varphi(x)| \leq m(\alpha_1, \alpha_2, \ldots, \alpha_n) \quad \forall x \in [a, b].$$

Luego, si $m(\alpha_1, \alpha_2, \ldots, \alpha_n)$ es pequeño, entonces $|f(x) - \varphi(x)|$ es todavía menor, para cada $x \in [a, b]$; por tanto, $\varphi(x)$ será más próximo a $f(x)$, en la medida que $m(\alpha_1, \alpha_2, \ldots, \alpha_n)$ sea menor; por ende, las gráficas de f y φ serán muy parecidas si $m(\alpha_1, \alpha_2, \ldots, \alpha_n)$ es muy pequeño. En términos de la jerga que hemos utilizado en este capítulo

$$m(\alpha_1, \alpha_2, \ldots, \alpha_n) = \|f - \varphi\|_\infty .$$

Luego, de entre todas las funciones $y = \varphi(x, \alpha_1, \alpha_2, \ldots, \alpha_n)$, se busca aquella para la cual $m(\alpha_1, \alpha_2, \ldots, \alpha_n)$ sea mínimo; es decir, se tiene que encontrar, de entre todos los conjuntos admisibles de parámetros $\alpha_1, \alpha_2, \ldots, \alpha_n$,

$$\underset{(\alpha_1,\alpha_2,\ldots,\alpha_n)}{\text{mín}} \left\| f - \varphi_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} \right\|_\infty$$

donde las $\varphi_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}$ son las curvas parametrizadas dadas por (4.60). Es así como Chebichev introduce la norma uniforme (también llamada norma de Chebichev) para resolver este tipo de problemas. Note que

$$\underset{(\alpha_1,\alpha_2,\ldots,\alpha_n)}{\text{mín}} \left\| f - \varphi_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} \right\|_\infty = \underset{(\alpha_1,\alpha_2,\ldots,\alpha_n)}{\text{mín}} \ \underset{a \leq x \leq b}{\text{máx}} |f(x) - \varphi(x)| ;$$

razón por la cual algunos métodos actuales de solución de problemas análogos a éste son llamados métodos “mínimax”.

Problema de la cuerda vibrante

Supongamos ahora que una cuerda de longitud L está sujeta en sus dos extremos, el izquierdo en el origen de coordenadas y el derecho en $(0, L)$. La cuerda se encuentra tensa con una fuerza de tensión

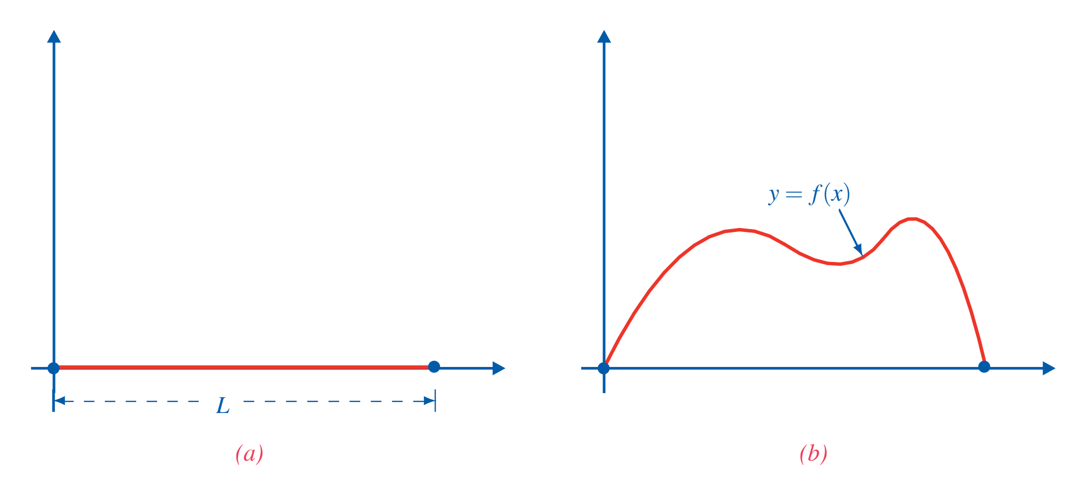


Figura 4-33 • *(a)* Cuerda tensa, sujeta en sus dos extremos, de longitud L en estado de equilibrio (sin vibrar). *(b)* "Instantánea", en el tiempo $t = \tau$, de la cuerda en uno de sus estados de oscilación.

T, como se muestra en la figura 4-33*(a)*. En el instante $t = 0$ se modifica su estado estirándola de su posición de equilibrio para que comience a vibrar en el plano x, y (se supone que la amplitud de la vibración es pequeña). En el instante $t = \tau$, la cuerda tiene la forma de una función $y = f(x)$ como se ilustra en la figura 4-33*(b)*. Se desea encontrar una función $y = \varphi(x)$ que aproxime la forma que tiene la cuerda en el instante $t = \tau$; es decir, que aproxime a la curva $y = f(x)$. El criterio de proximidad en este caso será físico más que geométrico. Para ello pensemos que tenemos otra cuerda exactamente igual que la primera, sujeta también en sus extremos en sendos puntos $(0,0)$ y $(0,L)$, pero cuya forma es precisamente la gráfica de la aproximación $y = \varphi(x)$ como se muestra en la figura 4-34. El criterio de proximidad que utilizaremos en este caso será mediante la energía potencial; es decir, la mejor aproximación será aquella para la cual la diferencia de energía potencial entre las cuerdas $y = f(x)$ y $y = \varphi(x)$ sea mínima. Calculemos la energía potencial para la cuerda $y = f(x)$; es decir, el trabajo realizado por la fuerza de tensión para obtener la forma de esta cuerda desde la posición de equilibrio. Sea x un punto en el intervalo $[0,L]$ y dx un pequeño incremento de x y calculemos el diferencial del trabajo, dU, realizado por la fuerza de tensión T para estirar el segmento de cuerda $[x, x+dx]$ de la posición de equilibrio a la forma de la curva $y = f(x)$ en este intervalo.

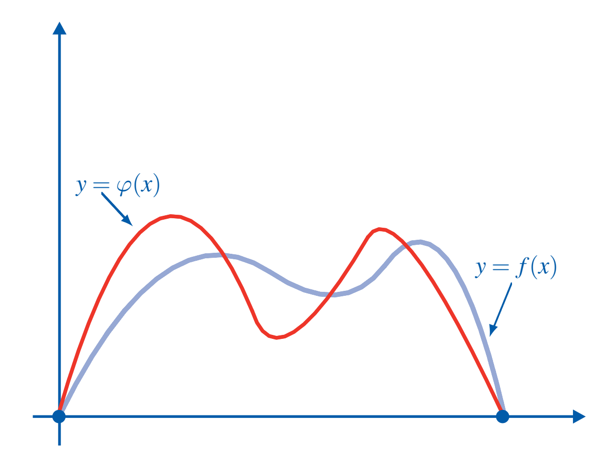


Figura 4-34 • La misma "instantánea" de la cuerda $y = f(x)$ en el instante $t = \tau$ y una aproximación a ella, $y = \varphi(x)$.

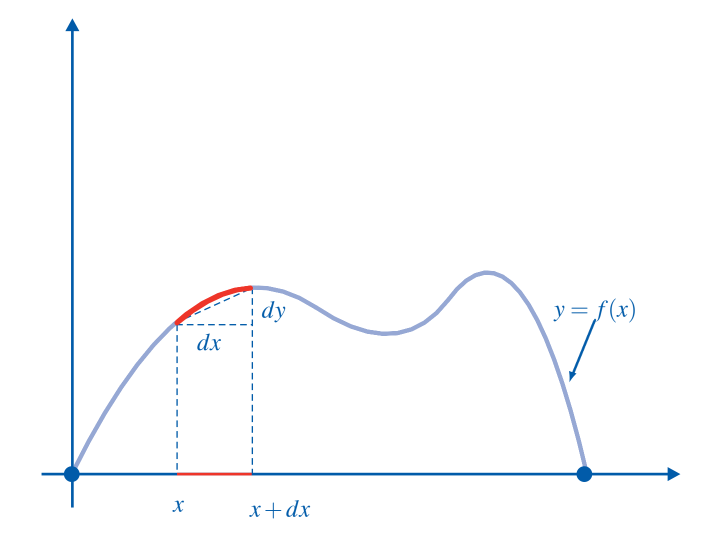


Figura 4-35 • La cuerda $y = f(x)$ en el instante $t = \tau$. La longitud del segmento de esta curva en el intervalo $[x, x + dx]$ se puede aproximar mediante la hipotenusa del triángulo rectángulo cuyos catetos son dx y dy.

De la figura 4-35 se desprende que el incremento de longitud en este intervalo es aproximadamente

$$\begin{aligned}\sqrt{(dx)^2 + (dy)^2} &= \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \\ &= \sqrt{1 + (f'(x))^2}\, dx\end{aligned}$$

Al multiplicar por la fuerza de tensión T la diferencia de longitudes entre la cuerda en el estado $y = f(x)$ y la cuerda en estado de equilibrio en el intervalo $[x, x + dx]$, obtenemos el diferencial de la energía potencial

$$dU = T\left(\sqrt{1 + (f'(x))^2}\, dx - dx\right);$$

esto es,

$$dU = T\left(\sqrt{1 + (f'(x))^2} - 1\right) dx.$$

Dado que la amplitud de las oscilaciones es pequeña, podemos utilizar el desarrollo de Taylor para obtener

$$\sqrt{1 + (f'(x))^2} \approx 1 + \frac{1}{2}(f'(x))^2.$$

Luego,

$$dU \approx \frac{T}{2}(f'(x))^2\, dx.$$

Por tanto, la energía potencial total de la cuerda es

$$U = \frac{T}{2}\int_0^L (f'(x))^2\, dx.$$

Así, la diferencia de energías potenciales entre las cuerdas $y = f(x)$ y $y = \varphi(x)$ está dada por

$$\frac{T}{2}\int_0^L (f'(x))^2 dx - \frac{T}{2}\int_0^L (\varphi'(x))^2 dx.$$

Ahora bien, pongamos para simplificar por el momento notación,

$$\alpha^2 = \int_0^L (f'(x))^2 dx \quad \text{y}$$
$$\beta^2 = \int_0^L (\varphi'(x))^2 dx.$$

Entonces,

$$\begin{aligned}|\alpha^2 - \beta^2| &= |(\alpha+\beta)(\alpha-\beta)| \\ &\le (|\alpha| + |\beta|)|\alpha - \beta|;\end{aligned}$$

esto es,

$$\left|\int_0^L (f'(x))^2 dx - \int_0^L (\varphi'(x))^2 dx\right| \le \left(\sqrt{\int_0^L (f'(x))^2 dx} + \sqrt{\int_0^L (\varphi'(x))^2 dx}\right) \cdot \left|\sqrt{\int_0^L (f'(x))^2 dx} - \sqrt{\int_0^L (\varphi'(x))^2 dx}\right|.$$

Pero $\sqrt{\int_0^L (g'(x))^2 dx} = \|g'\|$, la norma inducida por el producto interior $\langle g, h\rangle = \int_0^L g(x)h(x)dx$ (norma cuadrado medio); por tanto

$$\left|\int_0^L (f'(x))^2 dx - \int_0^L (\varphi'(x))^2 dx\right| \le \left(\sqrt{\int_0^L (f'(x))^2 dx} + \sqrt{\int_0^L (\varphi'(x))^2 dx}\right) \|f' - \varphi'\|.$$

Como hemos supuesto que las oscilaciones son pequeñas y que se trata de la misma cuerda tanto para $y = f(x)$ como para $y = \varphi(x)$, podemos suponer que existe $M > 0$ tal que $\int_0^L (f'(x))^2 dx \le M^2$ y $\int_0^L (\varphi'(x))^2 dx \le M^2$ (para cualquier aproximación φ). Entonces

$$\left|\int_0^L (f'(x))^2 dx - \int_0^L (\varphi'(x))^2 dx\right| \le 2M \|f' - \varphi'\|.$$

Por lo que si $\|f' - \varphi'\|$ es mínima, entonces la diferencia de energía potencial

$$\left|\int_0^L (f'(x))^2 dx - \int_0^L (\varphi'(x))^2 dx\right|$$

también será mínima. Así, eligiendo φ de tal suerte que $\|f' - \varphi'\|$ sea mínima, se habrá cumplido el criterio de proximidad.

De esta manera, hemos visto cómo surge, en problemas de aplicación, la norma cuadrado medio $\|g\| = \sqrt{\int_a^b (g(x))^2 dx}$ en $C[a, b]$ y también cómo se aplica para aproximar una función.

Comparación de las normas $\|g\| = \sqrt{\int_a^b (g(x))^2 dx}$ y $\|g\|_\infty$

Supongamos que se tiene una sucesión de funciones φ_n continuas en $[a,b]$ que converge uniformemente a $f \in C[a,b]$; i.e.,

$$\lim_{n\to\infty} \|f - \varphi_n\|_\infty = 0$$

entonces, dado que

$$\int_a^b |f(x) - \varphi_n(x)|^2 dx \leq \int_a^b \|f - \varphi_n\|_\infty^2 dx = (b-a)\|f - \varphi_n\|_\infty^2$$

se sigue que

$$\lim_{n\to\infty} \int_a^b |f(x) - \varphi_n(x)|^2 dx = 0;$$

es decir,

$$\lim_{n\to\infty} \|f - \varphi_n\| = 0 \tag{4.61}$$

Por tanto, convergencia uniforme implica convergencia en promedio cuadrático. Sin embargo, no es difícil ver que convergencia en promedio cuadrático no implica convergencia uniforme (cfr. el ejercicio resuelto 47, pág. 382). Esta es, aparentemente, una ventaja de la norma uniforme respecto a la norma cuadrado medio. Mas como explicamos arriba, la elección de la norma depende de la naturaleza del problema que se esté tratando y no de una elección arbitraria de ésta.[33] Pero, en realidad, la norma cuadrado medio tiene también una fortaleza muy importante que la convierte en una herramienta con un espectro de aplicación mayor que la norma uniforme. Aunque la convergencia en promedio cuadrático no implica convergencia uniforme, se puede probar que si la sucesión (φ_n) converge en promedio cuadrático a la función f, i.e., se cumple (4.61), entonces la sucesión (φ_n) contiene una subsucesión ψ_n que converge puntualmente a la función f, excepto en un conjunto de puntos del intervalo $[a,b]$ que tiene medida cero; esto es

$$\lim_{n\to\infty} \psi_n(x) = f(x)$$

para cada punto $x \in [a,b] - Z$ donde Z es un subconjunto de $[a,b]$ de "longitud cero"; por lo general, un conjunto finito o una sucesión de puntos del intervalo $[a,b]$ (un conjunto numerable). Así, la subsucesión ψ_n aproxima puntualmente a la función f en "casi todos los puntos" del intervalo $[a,b]$ y se puede mostrar que también converge en promedio a la función f. En muchos problemas, con esto es suficiente, ahí radica la importancia que tiene la norma cuadrado medio y la convergencia en promedio cuadrático.

❍ **Nota 4.19** Se puede probar que en realidad dicha subsucesión, (ψ_n), se puede elegir de tal suerte que dado cualquier $\varepsilon > 0$, existe un subconjunto M del intervalo $[a,b]$ de longitud inferior a ε, tal que la subsucesión (ψ_n) converge uniformemente a la función f en $[a,b] - M$ (y, por tanto, también en promedio).

[33] Recuerde que este no es el caso de espacios vectoriales de dimensión finita; pues en ellos todo par de normas son equivalentes.

4.3 Ejercicios resueltos y ejercicios propuestos

4.3.1 Ejercicios resueltos

Espacios con producto interior

1 En $\mathbb{R}^3$ si $\vec{x} = (x_1, x_2, x_3)$ y $\vec{y} = (y_1, y_2, y_3)$ se define

$$\langle \vec{x}, \vec{y} \rangle = y_1x_1 - 2y_1x_2 - 2y_2x_1 + 6y_2x_2 + y_2x_3 + y_3x_2 + y_3x_3.$$

Probar que $\langle \vec{x}, \vec{y} \rangle$ es un producto interior en $\mathbb{R}^3$.

DEMOSTRACIÓN ■ Sean $\vec{u} = (u_1, u_2, u_3), \vec{v} = (v_1, v_2, v_3), \vec{w} = (w_1, w_2, w_3) \in \mathbb{R}^3$ y $\lambda \in \mathbb{R}$:

1. (Simetría)

$$\langle \vec{u}, \vec{v} \rangle = v_1u_1 - 2v_1u_2 - 2v_2u_1 + 6v_2u_2 + v_2u_3 + v_3u_2 + v_3u_3$$

y

$$\begin{aligned} \langle \vec{v}, \vec{u} \rangle &= u_1v_1 - 2u_1v_2 - 2u_2v_1 + 6u_2v_2 + u_2v_3 + u_3v_2 + u_3v_3 \\ &= v_1u_1 - 2v_1u_2 - 2v_2u_1 + 6v_2u_2 + v_2u_3 + v_3u_2 + v_3u_3; \end{aligned}$$

es decir, $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$.

2. (Homogeneidad)

$$\begin{aligned} \langle \lambda\vec{u}, \vec{v} \rangle &= \langle (\lambda u_1, \lambda u_2, \lambda u_3), (v_1, v_2, v_3) \rangle \\ &= v_1(\lambda u_1) - 2v_1(\lambda u_2) - 2v_2(\lambda u_1) \\ &\quad + 6v_2(\lambda u_2) + v_2(\lambda u_3) + v_3(\lambda u_2) + v_3(\lambda u_3) \\ &= \lambda(v_1u_1 - 2v_1u_2 - 2v_2u_1 + 6v_2u_2 + v_2u_3 + v_3u_2 + v_3u_3) \\ &= \lambda \langle \vec{u}, \vec{v} \rangle. \end{aligned}$$

3. (Aditividad)

$$\begin{aligned} \langle \vec{u}, \vec{v} + \vec{w} \rangle &= \langle (u_1, u_2, u_3), (v_1 + w_1, v_2 + w_2, v_3 + w_3) \rangle \\ &= (v_1 + w_1)u_1 - 2(v_1 + w_1)u_2 - 2(v_2 + w_2)u_1 \\ &\quad + 6(v_2 + w_2)u_2 + (v_2 + w_2)u_3 + (v_3 + w_3)u_2 + (v_3 + w_3)u_3 \\ &= v_1u_1 - 2v_1u_2 - 2v_2u_1 + 6v_2u_2 + v_2u_3 + v_3u_2 + v_3u_3 \\ &\quad + w_1u_1 - 2w_1u_2 - 2w_2u_1 + 6w_2u_2 + w_2u_3 + w_3u_2 + w_3v_3 \\ &= \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle. \end{aligned}$$

4. (Positividad)

$$\begin{aligned} \langle \vec{u}, \vec{u} \rangle &= u_1^2 - 4u_1u_2 + 6u_2^2 + 2u_2u_3 + u_3^2 \\ &= u_1^2 - 4u_1u_2 + 4u_2^2 + u_2^2 + 2u_2u_3 + u_3^2 + u_2^2 \\ &= (u_1 - 2u_2)^2 + (u_2 + u_3)^2 + u_2^2 \\ &\geq 0. \end{aligned}$$

5. (Positividad)
$$\begin{aligned} \langle \vec{u}, \vec{u} \rangle &= 0 \Leftrightarrow \\ (u_1 - 2u_2)^2 + (u_2 + u_3)^2 + u_2^2 &= 0 \Leftrightarrow \\ \vec{u}_1 &= u_2 = u_3 = 0 \Leftrightarrow \\ \vec{u} &= \vec{0}_{\mathbb{R}^3}. \end{aligned}$$

De 1, 2, 3, 4 y 5, $\langle \vec{x}, \vec{y} \rangle$ es un producto interior. ■

En los ejercicios 2 a 4, **P** es el espacio de polinomios.

2 Probar que para todo $f \in \mathbf{P}$ la integral impropia
$$\int_0^\infty f(x)e^{-x}dx$$
converge.

DEMOSTRACIÓN ■ Sea f un polinomio; claramente si $f = \theta$ es el polinomio constante cero, la integral impropia converge al valor real cero. Si f tiene grado 0, f es un polinomio constante, $f(x) = k\ \ \forall x$; entonces
$$\begin{aligned} \int_0^\infty f(x)e^{-x}dx &= \lim_{r\to\infty} \int_0^r ke^{-x}dx \\ &= k \lim_{r\to\infty} \left[-e^{-x}\right]_{x=0}^{x=r} \\ &= k \lim_{r\to\infty} \left[-e^{-r} + 1\right] \\ &= k. \end{aligned}$$

Si f tiene grado $n = 1$, entonces $f(x) = a + bx$ para algún par de números reales a, b y
$$\begin{aligned} \int_0^\infty f(x)e^{-x}dx &= \lim_{r\to\infty} \int_0^r (a+bx)e^{-x}dx \\ &= \int_0^\infty ae^{-x}dx + b \lim_{r\to\infty} \int_0^r xe^{-x}dx \\ &= a + b \lim_{r\to\infty} \left[-xe^{-x}\Big|_{x=0}^{x=r} + \int_0^r e^{-x}dx\right] \\ &= a + b \lim_{r\to\infty} \left[-xe^{-x}\Big|_{x=0}^{x=r} - e^{-x}\Big|_{x=0}^{x=r}\right] \\ &= a + b \lim_{r\to\infty} \left[-\frac{r}{e^r} - (e^{-r} - 1)\right] \\ &= a + b; \end{aligned}$$

así que la integral impropia converge. Sea f un polinomio de grado n y suponga que la integral impropia $\int_0^\infty g(x)e^{-x}dx$ converge para todo polinomio de grado $n-1$. Entonces, ya que
$$\begin{aligned} \int_0^\infty f(x)e^{-x}dx &= \lim_{r\to\infty} \int_0^r f(x)e^{-x}dx \\ &= \lim_{r\to\infty} \left[-f(x)e^{-x}\Big|_{x=0}^{x=r} + \int_0^r f'(x)e^{-x}dx\right] \\ &= \lim_{r\to\infty} \left[f(0) - \frac{f(r)}{e^r}\right] + \lim_{r\to\infty} \int_0^r f'(x)e^{-x}dx \end{aligned}$$

y por hipótesis de inducción $\int_0^\infty f'(x)e^{-x}dx$ converge ($f'(x)$ es un polinomio de grado $n-1$) y

$$\lim_{r\to\infty}\frac{f(r)}{e^r}=0$$

(al aplicar sucesivamente la regla de L'Hôpital eventualmente se obtendrá el cociente de una constante sobre la función $r\mapsto e^r$ que tiende a cero conforme r tiende a infinito), se deduce que

$$\int_0^\infty f(x)e^{-x}dx$$

converge. ■

3 Deducir del ejercicio precedente que para cada par de polinomios $f,g\in\mathbf{P}$ la integral impropia

$$\int_0^\infty f(x)g(x)e^{-x}dx$$

converge.

Solución Si $f,g\in\mathbf{P}$, entonces fg es un polinomio también, por el ejercicio precedente $\int_0^\infty f(x)g(x)e^{-x}dx$ converge. ✔

4 Por el ejercicio anterior,

$$\langle f,g\rangle=\int_0^\infty f(x)g(x)e^{-x}dx$$

está definido para cada par de polinomios f y g. Probar que $\langle f,g\rangle$ es un producto interior en $\mathbf{P}$.

DEMOSTRACIÓN ■ Sean $f,g,h\in\mathbf{P}$ y $\lambda\in\mathbb{R}$:

(a)

$$\begin{aligned}\langle f,g\rangle &= \int_0^\infty f(x)g(x)e^{-x}dx\\ &= \lim_{r\to\infty}\int_0^\infty f(x)g(x)e^{-x}dx\\ &= \lim_{r\to\infty}\int_0^\infty g(x)f(x)e^{-x}dx\\ &= \int_0^\infty g(x)f(x)e^{-x}dx\\ &= \langle g,f\rangle .\end{aligned}$$

(b)

$$\begin{aligned}\langle \lambda f,g\rangle &= \int_0^\infty (\lambda f(x))g(x)e^{-x}dx\\ &= \lim_{r\to\infty}\int_0^r (\lambda f(x))g(x)e^{-x}dx\\ &= \lambda\lim_{r\to\infty}\int_0^r f(x)g(x)e^{-x}dx\\ &= \lambda\int_0^\infty f(x)g(x)e^{-x}dx\\ &= \lambda\langle f,g\rangle .\end{aligned}$$

(*c*)
$$\begin{aligned}\langle f, g+h\rangle &= \lim_{r\to\infty}\int_0^{\infty} f(x)(g(x)+h(x))e^{-x}dx\\ &= \lim_{r\to\infty}\left[\int_0^r f(x)g(x)e^{-x}dx+\int_0^r f(x)h(x)e^{-x}dx\right]\\ &= \lim_{r\to\infty}\int_0^r f(x)g(x)e^{-x}dx+\lim_{r\to\infty}\int_0^r f(x)h(x)e^{-x}dx\\ &= \int_0^{\infty} f(x)g(x)e^{-x}dx+\int_0^{\infty} f(x)h(x)e^{-x}dx\\ &= \langle f,g\rangle+\langle f,h\rangle\,.\end{aligned}$$

(*d*)
$$\langle f,f\rangle=\lim_{r\to\infty}\int_0^r (f(x))^2e^{-x}dx\geq 0.$$

(*e*) Claramente si $f=\theta$ es el polinomio constante cero, $\langle f,f\rangle=0$. Suponga que $f\in\mathbf{P}$ y $\langle f,f\rangle=0$, entonces

$$\lim_{r\to 0}\int_0^r (f(x))^2e^{-x}dx=0.$$

Si $f\neq\theta$ (la función constante cero), existe, por continuidad, un intervalo $[a,b]\subset[0,\infty)$ tal que $f^2(x)\neq 0$ para todo $x\in[a,b]$; luego

$$0<\int_a^b (f(x))^2e^{-x}dx\leq\int_0^{\infty}(f(x))^2e^{-x}dx=0$$

lo cual es imposible; por tanto, $f(x)=0$ para todo $x\in[0,\infty)$ y puesto que f es un polinomio, $f(x)=0$ para todo $x\in\mathbb{R}$; es decir, f es el polinomio constante cero (el único polinomio que tiene una infinidad de raíces es el polinomio constante cero).

De los 5 incisos precedentes se concluye que $\langle f,g\rangle$ es un producto interior. ■

En los ejercicios 5 a 9, $\langle f,g\rangle$ es el producto interior en el espacio de polinomios **P**, definido en el ejercicio 4 y $\|\cdot\|$ es la norma inducida por este producto.

5 Calcular $\langle 1+x,x\rangle$.

Solución

$$\begin{aligned}\langle 1+x,x\rangle &= \int_0^{\infty}(1+x)xe^{-x}dx\\ &= \lim_{r\to\infty}\left[\int_0^r xe^{-x}dx+\int_0^r x^2e^{-x}dx\right]\\ &= \lim_{r\to\infty}\left[\int_0^r xe^{-x}dx-x^2e^{-x}\big|_0^r+2\int_0^r xe^{-x}\right]\\ &= \lim_{r\to\infty}\left[3\left(-xe^{-x}\big|_0^r+\int_0^r e^{-x}\right)-x^2e^{-x}\big|_0^r\right]\\ &= \lim_{r\to\infty}\left[3\left(-xe^{-x}\big|_0^r-e^{-x}\big|_0^r\right)-x^2e^{-x}\big|_0^r\right]\\ &= \lim_{r\to\infty}\left[3\left(-re^{-r}-e^{-r}+1\right)-r^2e^{-r}\right]\\ &= 3\ \checkmark\end{aligned}$$

6 Calcular $\|x\|$.

Solución

$$\begin{aligned}
\|x\|^2 &= \langle x,x\rangle \\
&= \int_0^\infty x^2e^{-x}dx \\
&= \lim_{r\to\infty}\int_0^r x^2e^{-x}dx \\
&= \lim_{r\to\infty}\left[-x^2e^{-x}\Big|_0^r + 2\int_0^r xe^{-x}dx\right] \\
&= \lim_{r\to\infty}\left[-x^2e^{-x}\Big|_0^r + 2\left(-xe^{-x}\Big|_0^r + \int_0^r e^{-x}dx\right)\right] \\
&= \lim_{r\to\infty}\left[-x^2e^{-x}\Big|_0^r + 2\left(-xe^{-x}\Big|_0^r + e^{-x}\Big|_0^r\right)\right] \\
&= \lim_{r\to\infty}\left[-r^2e^{-r} + 2\left(-re^{-r} - e^{-r} + 1\right)\right] \\
&= 2.
\end{aligned}$$

Por tanto $\|x\| = \sqrt{2}$. ✔

7 Encontrar el ángulo ϕ entre $f(x) = 1$ y $g(x) = x$.

Solución

$$\begin{aligned}
\|1\|^2 &= \int_0^\infty e^{-x}dx \\
&= \lim_{r\to\infty}\int_0^r e^{-x}dx \\
&= \lim_{r\to\infty} -e^{-x}\Big|_0^r \\
&= \lim_{r\to\infty}\left[-e^{-r}+1\right] \\
&= 1.
\end{aligned}$$

Por el ejercicio 6 $\|x\| = \sqrt{2}$, y ya que

$$\begin{aligned}
\langle 1,x\rangle &= \int_0^\infty xe^{-x}dx \\
&= \lim_{r\to\infty}\int_0^r xe^{-x}dx \\
&= \lim_{r\to\infty}\left[-xe^{-x}\Big|_0^r - e^{-x}\Big|_0^r\right] \\
&= 1
\end{aligned}$$

se tiene

$$\begin{aligned}
\phi &= \arccos\left(\frac{\langle 1,x\rangle}{\|1\|\,\|x\|}\right) \\
&= \arccos\left(\frac{1}{\sqrt{2}}\right) \\
&= \arccos\left(\frac{\sqrt{2}}{2}\right) \\
&= \frac{\pi}{4}.
\end{aligned}$$ ✔

8 Hallar una base ortonormal del subespacio $S = \text{gn}(1, x, x^2)$.

Solución

$$\begin{aligned}\|1\|^2 &= \int_0^\infty e^{-x}dx \\ &= \lim_{r\to\infty}\int_0^r e^{-x}dx \\ &= \lim_{r\to\infty}\left[-e^{-x}\right]_0^r \\ &= \lim\left[1-e^{-r}\right] \\ &= 1,\end{aligned}$$

por tanto, $u_1 = 1$

$$\begin{aligned}w_2 &= x - \langle x, 1\rangle\, 1 \\ &= x - \int_0^\infty xe^{-x}dx \\ &= x-1,\end{aligned}$$

$$\begin{aligned}\|w_2\|^2 &= \int_0^\infty (x-1)^2 e^{-x}dx \\ &= 1\end{aligned}$$

por lo que,

$$u_2 = x - 1.$$

$$\begin{aligned}w_3 &= x^2 - \left\langle x^2, 1\right\rangle 1 - \left\langle x^2, x-1\right\rangle (x-1) \\ &= x^2 - \int_0^\infty x^2e^{-x}dx - \int_0^\infty x^2(x-1)e^{-x}dx(x-1) \\ &= x^2 - 2 - 4(x-1) \\ &= x^2 - 4x + 2,\end{aligned}$$

$$\begin{aligned}\|w_3\|^2 &= \int_0^\infty \left(x^2-4x+2\right)^2 e^{-x}dx \\ &= 4\end{aligned}$$

y, por tanto,

$$u_3 = \frac{1}{2}\left(x^2 - 4x + 2\right).$$

La base ortonormal es entonces

$$\left\{1, x-1, \tfrac{x^2}{2} - 2x + 1\right\}. \quad ✔$$

9 Encontrar la aproximación óptima para $f(x) = x^3$ en el subespacio $S = \text{gn}(1, x, x^2)$.

Solución Por el ejercicio precedente $\left\{1, x-1, \frac{x^2}{2} - 2x + 1\right\}$ es una base ortonormal para el subespacio S; por el teorema 4.15 el vector proyección p de f sobre S, es la aproximación óptima de este subespacio sobre f y por 4.14 del teorema 4.11 (cfr. pág. 269), el vector proyección está dado por:

$$p = \sum_{i=1}^{3} \langle f, u_i \rangle u_i.$$

$$\langle x^3, 1 \rangle = \int_0^\infty x^3 e^{-x} dx = 6,$$

$$\langle x^3, x-1 \rangle = \int_0^\infty x^3 (x-1) e^{-x} dx = 18,$$

$$\left\langle x^3, \tfrac{x^2}{2} - 2x + 1 \right\rangle = \int_0^\infty x^3 \left(\tfrac{x^2}{2} - 2x + 1 \right) e^{-x} dx = 18.$$

Por tanto,

$$\begin{aligned} p &= 6 + 18(x-1) + 18(\tfrac{x^2}{2} - 2x + 1) \\ &= 6 - 18x + 9x^2. \quad ✔ \end{aligned}$$

10 (**Representación matricial de un producto interior**). Sea $\langle \vec{u}, \vec{v} \rangle$ un producto interior definido en $\mathbb{R}^n$. Probar que existe una matriz simétrica $A \in \mathfrak{M}_n$ tal que

$$\langle \vec{u}, \vec{v} \rangle = \vec{u}^t A \vec{v}$$

para todo par de vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$. Se dice entonces que la matriz A es una representación matricial del producto interior $\langle \vec{u}, \vec{v} \rangle$.

DEMOSTRACIÓN ■ Sean $\vec{e}_i$, $i = 1, \ldots, n$ los vectores que forman la base canónica del espacio $\mathbb{R}^n$ y $A = [\langle \vec{e}_i, \vec{e}_j \rangle]$. Si $\vec{u} = (\vec{u}_1, \ldots, u_n)$ y $\vec{v} = (v_1, \ldots, v_n)$ son un par vectores en $\mathbb{R}^n$, entonces

$$\begin{aligned} \langle \vec{u}, \vec{v} \rangle &= \left\langle \sum_{i=1}^{n} u_i \vec{e}_i, \sum_{i=1}^{n} v_i \vec{e}_i \right\rangle \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} u_i \langle \vec{e}_i, \vec{e}_j \rangle v_j \\ &= \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \sum_{j=1}^{n} \langle \vec{e}_1, \vec{e}_j \rangle v_j \\ \vdots \\ \sum_{j=1}^{n} \langle \vec{e}_n, \vec{e}_j \rangle v_j \end{bmatrix} \\ &= \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \langle \vec{e}_1, \vec{e}_1 \rangle & \cdots & \langle \vec{e}_1, \vec{e}_n \rangle \\ \vdots & \ddots & \vdots \\ \langle \vec{e}_n, \vec{e}_1 \rangle & \cdots & \langle \vec{e}_n, \vec{e}_n \rangle \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\ &= \vec{u}^t A \vec{v}. \quad ■ \end{aligned}$$

11 Utilizar el ejercicio 10 para encontrar una representación matricial del producto interior definido en el ejercicio resuelto 1 de este apartado; es decir, una matriz simétrica $A \in \mathfrak{M}_3$ tal que $\langle \vec{u}, \vec{v} \rangle = \vec{u}^t A \vec{v}$.

Solución En este caso el producto interior está definido para cada par de vectores $\vec{x} = (x_1, x_2, x_3)$ y $\vec{y} = (y_1, y_2, y_3)$, por

$$\langle \vec{x}, \vec{y} \rangle = y_1 x_1 - 2y_1 x_2 - 2y_2 x_1 + 6y_2 x_2 + y_2 x_3 + y_3 x_2 + y_3 x_3,$$

entonces,

$$\begin{aligned}\langle\vec{e}_1,\vec{e}_1\rangle &= \langle(1,0,0),(1,0,0)\rangle = 1,\\ \langle\vec{e}_1,\vec{e}_2\rangle &= \langle(1,0,0),(0,1,0)\rangle = -2,\\ \langle\vec{e}_1,\vec{e}_3\rangle &= \langle(1,0,0),(0,0,1)\rangle = 0,\\ \langle\vec{e}_2,\vec{e}_2\rangle &= \langle(0,1,0),(0,1,0)\rangle = 6,\\ \langle\vec{e}_2,\vec{e}_3\rangle &= \langle(0,1,0),(0,0,1)\rangle = 1,\\ \langle\vec{e}_3,\vec{e}_3\rangle &= \langle(0,0,1),(0,0,1)\rangle = 1,\end{aligned}$$

por lo que

$$A = \begin{bmatrix} 1 & -2 & 0\\ -2 & 6 & 1\\ 0 & 1 & 1\end{bmatrix}.$$

Comprobación:

$$\begin{aligned}\begin{bmatrix} x_1 & x_2 & x_3\end{bmatrix}\begin{bmatrix} 1 & -2 & 0\\ -2 & 6 & 1\\ 0 & 1 & 1\end{bmatrix}\begin{bmatrix} y_1\\ y_2\\ y_3\end{bmatrix} &= y_1x_1 - 2y_1x_2 - 2y_2x_1 + 6x_2y_2\\ &\quad + y_2x_3 + y_3x_2 + y_3x_3\\ &= \langle\vec{x},\vec{y}\rangle. \quad ✔\end{aligned}$$

En los ejercicios 12 a 14, **E** es el conjunto de funciones continuas en el intervalo $[a,\infty)$ tales que la integral impropia

$$\int_a^\infty (f(x))^2 e^{-x^2}\,dx$$

converge.

12 Mostrar que si $f,g\in\mathbf{E}$, entonces $f\,g\in\mathbf{E}$. Probar, de hecho, que la integral impropia $\int_a^\infty f(x)\,g(x)\,e^{-x^2}\,dx$ converge absolutamente; i.e., que la integral impropia $\int_a^\infty |f(x)g(x)|\,e^{-x^2}\,dx$ converge.

DEMOSTRACIÓN ■ Dado que $|f|^2 = f^2$, la implicación $f\in\mathbf{E}\Rightarrow|f|\in\mathbf{E}$ es válida. Sean $f,g\in\mathbf{E}$ y $r>a$. Por la desigualdad de Schwarz, aplicada al producto interior $\left\langle |f(x)|\,e^{-x^2/2}, |g(x)|\,e^{-x^2/2}\right\rangle$ en $C[a,r]$, se tiene

$$\int_a^r |f(x)g(x)|\,e^{-x^2} \le \left(\int_a^r (f(x))^2 e^{-x^2}\,dx\right)^{1/2}\left(\int_a^r (g(x))^2 e^{-x^2}\,dx\right)^{1/2}.$$

Puesto que $f,g\in\mathbf{E}$, las integrales impropias

$$\int_a^\infty (f(x))^2 e^{-x^2}\,dx \quad \text{y} \quad \int_a^\infty (f(x))^2 e^{-x^2}\,dx$$

convergen; entonces, para todo $r>a$,

$$\int_a^r |f(x)g(x)|\,e^{-x^2} \le \left(\int_a^\infty (f(x))^2 e^{-x^2}\,dx\right)^{1/2}\left(\int_a^\infty (f(x))^2 e^{-x^2}\,dx\right)^{1/2};$$

así que la función creciente

$$r\mapsto \int_a^r |f(x)g(x)|\,e^{-x^2}$$

es acotada y, por tanto,

$$\int_a^\infty |f(x)g(x)|\,e^{-x^2}dx = \lim_{r\to\infty}\int_a^r |f(x)g(x)|\,e^{-x^2}$$

existe. Dado que la convergencia de $\int_a^\infty |u(x)|\,dx$ implica la convergencia de $\int_a^\infty u(x)dx$, se tiene que $\int_a^\infty f(x)g(x)e^{-x^2}dx$ converge. ■

13 Demostrar que **E**, con las operaciones usuales de suma entre funciones y multiplicación de un escalar con una función, es un espacio vectorial.

DEMOSTRACIÓN ■ Claramente la función constante cero en el intervalo $[a,\infty)$ pertenece a **E**. Puesto que la suma de funciones continuas y la multiplicación de un escalar con una función continua también es una función continua y $\mathbf{E}\subset\mathscr{F}([a,\infty))$, el espacio de funciones con dominio el intervalo $[a,\infty)$, basta probar que $f,g\in\mathbf{E}$ y $\alpha,\beta\in\mathbb{R}$ implican $\alpha f+\beta g\in\mathbf{E}$. En efecto:

$$\lim_{r\to\infty}\int_a^r(\alpha f(x)+\beta g(x))^2e^{-x^2}dx = \lim_{r\to\infty}\begin{bmatrix}\alpha^2\int_a^r(f(x))^2e^{-x^2}dx\\ +2\alpha\beta\int_a^r f(x)g(x)e^{-x^2}dx\\ +\beta^2\int_a^r(g(x))^2e^{-x^2}dx\end{bmatrix}$$

y ya que $f,g\in\mathbf{E}$, por definición de **E** y por el ejercicio precedente, $\int_a^\infty(f(x))^2e^{-x^2}dx$, $\int_a^\infty f(x)g(x)e^{-x^2}dx$, $\int_a^\infty(g(x))^2e^{-x^2}dx$ convergen y, por tanto, $\int_a^\infty(\alpha f(x)+\beta g(x))^2e^{-x^2}dx$ converge; es decir, $\alpha f+\beta g\in\mathbf{E}$. ■

14 Por el ejercicio 12, la integral impropia $\int_a^\infty(f(x)g(x))e^{-x^2}dx$ converge para todo par de funciones en el espacio **E** del ejercicio anterior, si se define $\langle f,g\rangle=\int_a^\infty f(x)g(x)e^{-x^2}dx$ probar que éste es un producto interior en **E**.

DEMOSTRACIÓN ■ Sean $f,g,h\in\mathbf{E}$ y $\alpha\in\mathbb{R}$.

(a)
$$\begin{aligned}\langle f,g\rangle &= \int_a^\infty f(x)g(x)e^{-x^2}dx\\ &= \int_a^\infty g(x)f(x)e^{-x^2}dx\\ &= \langle g,f\rangle.\end{aligned}$$

(b)
$$\begin{aligned}\langle \alpha f,g\rangle &= \int_a^\infty \alpha f(x)g(x)e^{-x^2}dx\\ &= \alpha\int_a^\infty f(x)g(x)e^{-x^2}dx\\ &= \alpha\langle f,g\rangle.\end{aligned}$$

(c)
$$\begin{aligned}\langle f,g+h\rangle &= \int_a^\infty f(x)(g(x)+h(x))e^{-x^2}dx\\ &= \int_a^\infty f(x)g(x)e^{-x^2}dx+\int_a^\infty f(x)h(x)e^{-x^2}dx\\ &= \langle f,g\rangle+\langle f,h\rangle.\end{aligned}$$

(d)
$$\langle f,f\rangle = \int_a^\infty f^2(x)e^{-x^2}dx \geq 0.$$

(e) Claramente si θ es la función constante cero en $[a,\infty)$, $\langle\theta,\theta\rangle = 0$. Sean $f \in \mathbf{E}$ tal que $\langle f,f\rangle = 0$ y $r > 0$ un número real fijo. Entonces,

$$0 \leq \int_a^r (f(x))^2 e^{-x^2}dx \leq \int_a^\infty (f(x))^2 e^{-x^2}dx = 0;$$

de donde, para la norma $\|u\| = \left(\int_a^r (u(x))^2dx\right)^{1/2}$ en $C[a,r]$, se tiene

$$\left\| f(x)e^{-x^2/2} \right\| = 0$$

y, por tanto,

$$f(x)e^{-x^2/2} = 0 \;\forall x \in [a,r]$$

y ya que $e^{-x^2/2} \neq 0$ para todo $x \in [a,r]$, se concluye que f es la función constante cero en el intervalo $[a,r]$. Dado que r es cualquier número real mayor que a, entonces $f(x) = 0$ para todo $x \in [a,\infty)$; es decir, $f = \theta$, el neutro aditivo de $\mathbf{E}$. ■

En los ejercicios 15 a 18, $\langle f,g\rangle = \int_a^\infty f(x)g(x)e^{-x^2}dx$ es el producto interior definido en el ejercicio 14, $\mathbf{E}$ es el espacio del ejercicio 13 y $\|f\| = \left(\int_a^\infty (f(x))^2 e^{-x^2}dx\right)^{1/2}$ la norma inducida por este producto escalar, para el caso particular $a = 0$.

15 Probar que $f \in \mathbf{E}$ y calcular $\|f\|$ si $f(x) = 1$.

Solución
$$\|f\|^2 = \int_0^\infty 1^2 e^{-x^2}dx = \int_0^\infty e^{-x^2}dx.$$

Sean $R > 0$, Q_R la región de los puntos $(x,y) \in \mathbb{R}^2$ con $0 \leq x \leq R$, $0 \leq y \leq R$ y C_R la región de puntos $(x,y) \in \mathbb{R}^2$ con $x^2 + y^2 \leq R^2$ en el primer cuadrante. Entonces

$$\iint_{C_R} e^{-x^2-y^2}dA \leq \iint_{Q_R} e^{-x^2-y^2}dA \leq \iint_{C_{\sqrt{2}R}} e^{-x^2-y^2}dA.$$

Ya que,

$$\iint_{Q_R} e^{-x^2-y^2}dA = \int_0^R \int_0^R e^{-x^2-y^2}dydx = \left(\int_0^R e^{-x^2}dx\right)^2$$

y, al emplear coordenadas polares,

$$\begin{aligned}
\iint_{C_\rho} e^{-x^2-y^2}dA &= \int_0^{\frac{\pi}{2}} \int_0^\rho e^{-r^2} r dr d\theta \\
&= \int_0^{\frac{\pi}{2}} \left[-\frac{1}{2}e^{-r^2}\right]_0^\rho d\theta \\
&= \frac{\pi}{4}\left[1 - e^{-\rho^2}\right]
\end{aligned}$$

se tiene

$$\frac{\pi}{4}\left[1-e^{-\frac{1}{2}R^2}\right] \le \left(\int_0^R e^{-x^2}dx\right)^2 \le \frac{\pi}{4}\left[1-e^{-R^2}\right]$$

y, por tanto,

$$\lim_{R\to\infty}\frac{\pi}{4}\left[1-e^{-\frac{1}{2}R^2}\right] \le \lim_{R\to\infty}\left(\int_0^R e^{-x^2}dx\right)^2 \le \lim_{R\to\infty}\frac{\pi}{4}\left[1-e^{-R^2}\right];$$

de donde

$$\lim_{R\to\infty}\left(\int_0^R e^{-x^2}dx\right)^2 = \frac{\pi}{4}.$$

Luego,

$$\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2};$$

lo cual implica $1 \in \mathbf{E}$ y

$$\|1\| = \left(\frac{\sqrt{\pi}}{2}\right)^{1/2} = \frac{\sqrt[4]{\pi}}{\sqrt{2}}. \quad ✔$$

16 Probar que la función $f(x)=\sqrt{x}$, $x\ge 0$, pertenece a $\mathbf{E}$ y calcular $\|f\|$.

DEMOSTRACIÓN ■

$$\begin{aligned}\int_0^\infty (\sqrt{x})^2 e^{-x^2}dx &= \lim_{r\to\infty}\int_0^r xe^{-x^2}dx\\ &= \lim_{r\to\infty}\left[-\frac{1}{2}e^{-x^2}\right]_0^r\\ &= \lim_{r\to\infty}\left[\frac{1}{2}-\frac{1}{2}e^{-r^2}\right]\\ &= \frac{1}{2}.\end{aligned}$$

Lo cual prueba que $f\in\mathbf{E}$. Por último,

$$\begin{aligned}\|f\| &= \left(\int_0^\infty (\sqrt{x})^2 e^{-x^2}dx\right)^{1/2}\\ &= \frac{\sqrt{2}}{2}. \quad ■\end{aligned}$$

17 Demostrar que $f\in\mathbf{E}$ y calcular $\|f\|$ si $f(x)=x$.

DEMOSTRACIÓN ■

$$\begin{aligned}\int_0^r x^2e^{-x^2}dx &= \int_0^r xe^{-x^2}x\,dx\\ &= -\frac{1}{2}xe^{-x^2}\Big|_0^r + \frac{1}{2}\int_0^r e^{-x^2}dx.\end{aligned}$$

Luego,

$$\begin{aligned}\int_0^{\infty} x^2e^{-x^2}dx &= \lim_{r\to\infty}\int_0^r x^2e^{-x^2}dx\\ &= \lim_{r\to\infty}\left[-\frac{1}{2}xe^{-x^2}\Big|_0^r + \frac{1}{2}\int_0^r e^{-x^2}dx\right]\\ &= \lim_{r\to\infty}\left[-\frac{1}{2}re^{-r^2}\right] + \frac{1}{2}\lim_{r\to\infty}\int_0^r e^{-x^2}dx\\ &= 0+\frac{\sqrt{\pi}}{4}\\ &= \frac{\sqrt{\pi}}{4}.\end{aligned}$$

Por lo que $f\in\mathbf{E}$ y

$$\|f\| = \left(\frac{\sqrt{\pi}}{4}\right)^{1/2} = \frac{\sqrt[4]{\pi}}{2}. \quad ■$$

18 Sean $\alpha\in\mathbb{R}$, $f(x)=\cos(\alpha x)$. Probar que $f\in\mathbf{E}$ y calcular $\langle 1,f\rangle$ (ya se probó en el ejercicio 15 que la función constante 1 pertenece a $\mathbf{E}$).

DEMOSTRACIÓN ■ Puesto que $\cos^2(\alpha x)e^{-x^2}\le e^{-x^2}$ para todo x y $\int_0^{\infty} e^{-x^2}dx$ converge (cfr. ejercicio 15), se concluye que $\int_0^{\infty}\cos^2(\alpha x)e^{-x^2}dx$ converge y, por tanto, $f\in\mathbf{E}$ para todo $\alpha\in\mathbb{R}$. Sea

$$\varphi(\alpha) = \int_0^{\infty} e^{-x^2}\cos(\alpha x)dx.$$

Entonces, puesto que la función $(x,\alpha)\mapsto e^{-x^2}\cos(\alpha x)$ y su derivada parcial $\partial/\partial\alpha(\,e^{-x^2}\cos(\alpha x)) = -xe^{-x^2}\operatorname{sen}(\alpha x)$ son continuas,

$$\begin{aligned}\left|e^{-x^2}\cos(\alpha x)\right| &\le e^{-x^2},\\ \left|-xe^{-x^2}\operatorname{sen}(\alpha x)\right| &\le |x|\,e^{-x^2}\end{aligned}$$

y las integrales impropias $\int_0^{\infty} e^{-x^2}dx$, $\int_0^{\infty}|x|\,e^{-x^2}dx$ convergen (por los ejercicios 15 y 16), se tiene

$$\begin{aligned}\varphi'(\alpha) &= \int_0^{\infty}\frac{\partial}{\partial\alpha}\left(e^{-x^2}\cos(\alpha x)\right)dx\\ &= \int_0^{\infty} -xe^{-x^2}\operatorname{sen}(\alpha x)dx.\end{aligned}$$

Al integrar por partes,

$$\begin{aligned}\varphi'(\alpha) &= \lim_{r\to\infty}\left[\frac{1}{2}e^{-x^2}\operatorname{sen}(\alpha x)\right]_{x=0}^{x=r} - \frac{\alpha}{2}\int_0^{\infty} e^{-x^2}\cos(\alpha x)dx\\ &= \lim_{r\to\infty}\left[\frac{1}{2}e^{-r^2}\operatorname{sen}(\alpha r)\right] - \frac{\alpha}{2}\varphi(\alpha).\end{aligned}$$

Y puesto que

$$0\le\left|\frac{1}{2}e^{-r^2}\operatorname{sen}(\alpha r)\right|\le\frac{1}{2}e^{-r^2}$$

se desprende que

$$\lim_{r\to 0}\left[\frac{1}{2}e^{-r^2}\operatorname{sen}(\alpha r)\right]=0$$

y, por tanto,

$$\varphi'(\alpha)=-\frac{\alpha}{2}\varphi(\alpha);$$

es decir,

$$\frac{d\varphi}{d\alpha}=-\frac{\alpha}{2}\varphi.$$

Al separar variables se obtiene

$$\frac{d\varphi}{\varphi}=-\frac{1}{2}\alpha d\alpha,$$

y al integrar

$$\ln|\varphi|=-\frac{1}{4}\alpha^2+C$$

para alguna constante C. Esto es,

$$\varphi(\alpha)=Ae^{-\frac{\alpha^2}{4}}$$

para alguna constante A. Por el ejercicio 15 se sabe que

$$\varphi(0)=\int_0^\infty e^{-x^2}dx=\frac{\sqrt{\pi}}{2}$$

y, por tanto, $A=\frac{\sqrt{\pi}}{2}$. Luego

$$\begin{aligned}\langle 1,f\rangle &= \langle 1,\cos(\alpha x)\rangle\\ &= \int_0^\infty e^{-x^2}\cos(\alpha x)dx\\ &= \frac{\sqrt{\pi}}{2}e^{-\alpha^2/4}. \quad ■\end{aligned}$$

19 Calcular el ángulo entre las funciones $f(x)=1$ y $g(x)=x$.

Solución De los ejercicios 16, 17 y 15

$$\begin{aligned}\langle 1,x\rangle &= \int_0^\infty xe^{-x^2}dx=\frac{1}{2},\\ \|x\| &= \frac{\sqrt[4]{\pi}}{2} \quad \text{y}\\ \|1\| &= \frac{\sqrt[4]{\pi}}{\sqrt{2}}.\end{aligned}$$

Por lo que,

$$\begin{aligned}\phi &= \arccos\left[\frac{\langle 1,x\rangle}{\|1\|\,\|x\|}\right]\\ &= \arccos\left[\frac{\frac{1}{2}}{\frac{\sqrt[4]{\pi}}{\sqrt{2}}\frac{\sqrt[4]{\pi}}{\sqrt{2}}}\right]\end{aligned}$$

$$= \arccos\left[\frac{1}{\sqrt{\pi}}\right]$$
$$\approx 0.97134$$

o

$$\phi \approx 55.65°. \quad ✔$$

20 Encontrar una base ortonormal para el subespacio $S = \text{gn}(1, x)$ utilizando el proceso de ortogonalización de Gram-Schmidt.

Solución Por los ejercicios 16, 17 y 15

$$u_1 = \frac{1}{\|1\|} = \frac{1}{\frac{\sqrt[4]{\pi}}{\sqrt{2}}} = \frac{\sqrt{2}}{\sqrt[4]{\pi}}$$

$$w_2 = v_2 - \langle v_2, u_1 \rangle u_1$$
$$= x - \left\langle x, \frac{\sqrt{2}}{\sqrt[4]{\pi}} \right\rangle \frac{\sqrt{2}}{\sqrt[4]{\pi}}.$$

$$\left\langle x, \frac{\sqrt{2}}{\sqrt[4]{\pi}} \right\rangle = \frac{\sqrt{2}}{\sqrt[4]{\pi}} \int_0^\infty x e^{-x^2} dx$$
$$= \frac{1}{2\sqrt[4]{\pi}}\sqrt{2},$$
$$\left\langle x, \frac{\sqrt{2}}{\sqrt[4]{\pi}} \right\rangle \frac{\sqrt{2}}{\sqrt[4]{\pi}} = \frac{1}{2\sqrt[4]{\pi}}\sqrt{2}\frac{\sqrt{2}}{\sqrt[4]{\pi}}$$
$$= \frac{1}{\sqrt{\pi}}$$

y, por tanto,

$$w_2 = x - \left\langle x, \frac{\sqrt{2}}{\sqrt[4]{\pi}} \right\rangle \frac{\sqrt{2}}{\sqrt[4]{\pi}} = x - \frac{1}{\sqrt{\pi}}.$$

$$\|w_2\|^2 = \int_0^\infty e^{-x^2} (x - \frac{1}{\sqrt{\pi}})^2 dx$$
$$= \int_0^\infty e^{-x^2} (x^2 - 2\frac{x}{\sqrt{\pi}} + \frac{1}{\pi}) dx$$
$$= \frac{1}{4}\frac{-2+\pi}{\sqrt{\pi}}.$$

Así que,

$$u_2 = \frac{2\sqrt[4]{\pi}}{\sqrt{\pi - 2}}(x - \frac{1}{\sqrt{\pi}}).$$

La base ortonormal es, entonces,

$$\left\{\frac{\sqrt{2}}{\sqrt[4]{\pi}}, \frac{2\sqrt[4]{\pi}}{\sqrt{\pi - 2}}(x - \frac{1}{\sqrt{\pi}})\right\}. \quad ✔$$

21 Encontrar la aproximación óptima para $u(x) = x^2$ en el subespacio $S = \text{gn}(1,x)$.

Solución La aproximación óptima para u en el subespacio S viene dada por la proyección ortogonal p de u sobre S. Y ya que por el ejercicio precedente $u_1 = \frac{\sqrt{2}}{\sqrt[4]{\pi}}$ y $u_2 = \frac{2\sqrt[4]{\pi}}{\sqrt{\pi-2}}(x - \frac{1}{\sqrt{\pi}})$ forman una base ortonormal de S, por (4.14) del teorema 4.11 (cfr. pág. 269)

$$p = \langle u, u_1 \rangle u_1 + \langle u, u_2 \rangle u_2.$$

Del ejercicio 17,

$$\begin{aligned}\langle u, u_1 \rangle u_1 &= \left(\tfrac{\sqrt{2}}{\sqrt[4]{\pi}}\right)^2 \int_0^\infty x^2 e^{-x^2} dx \\ &= \tfrac{2}{\sqrt{\pi}} \tfrac{\sqrt{\pi}}{4} \\ &= \tfrac{1}{2},\end{aligned}$$

$$\begin{aligned}\langle u, u_2 \rangle &= \tfrac{2\sqrt[4]{\pi}}{\sqrt{\pi-2}} \left[\int_0^\infty x^3 e^{-x^2} dx - \tfrac{1}{\sqrt{\pi}} \int_0^\infty x^2 e^{-x^2} dx\right] \\ &= \tfrac{2\sqrt[4]{\pi}}{\sqrt{\pi-2}} \left[\int_0^\infty x^3 e^{-x^2} dx - \tfrac{1}{\sqrt{\pi}} \tfrac{\sqrt{\pi}}{4}\right]\end{aligned}$$

$$\begin{aligned}\int_0^\infty x^3 e^{-x^2} dx &= \int_0^\infty x^2 x e^{-x^2} dx \\ &= \lim_{r\to\infty} \left[-\tfrac{1}{2} x^2 e^{-x^2}\Big|_0^r + \tfrac{1}{2}\int_0^r 2x e^{-x^2} dx\right] \\ &= \lim_{r\to\infty} \left[-\tfrac{1}{2} r^2 e^{-r^2} - \tfrac{1}{2}\left[e^{-x^2}\right]_0^r\right] \\ &= \tfrac{1}{2}.\end{aligned}$$

Por lo que,

$$\begin{aligned}\langle u, u_2 \rangle &= \tfrac{2\sqrt[4]{\pi}}{\sqrt{\pi-2}} \left[\tfrac{1}{2} - \tfrac{1}{4}\right] \\ &= \tfrac{\sqrt[4]{\pi}}{2\sqrt{\pi-2}}\end{aligned}$$

y

$$\begin{aligned}\langle u, u_2 \rangle u_2 &= \tfrac{\sqrt[4]{\pi}}{2\sqrt{\pi-2}} \tfrac{2\sqrt[4]{\pi}}{\sqrt{\pi-2}} (x - \tfrac{1}{\sqrt{\pi}}) \\ &= \tfrac{\sqrt{\pi}}{\pi-2}\left(x - \tfrac{1}{\sqrt{\pi}}\right).\end{aligned}$$

Por tanto, la aproximación óptima de $u(x) = x^2$ en el espacio $S = \text{gn}(1,x)$ es

$$p = \frac{1}{2} + \frac{\sqrt{\pi}}{\pi-2}\left(x - \frac{1}{\sqrt{\pi}}\right). \quad ✔$$

22 Encontrar los valores de las constantes a, b, c tales que

$$\int_{-1}^{1} \left(e^x - a - bx - cx^2\right)^2 dx$$

alcance el valor más pequeño posible.

Solución El problema se resuelve si $a+bx+cx^2$ es la aproximación óptima a la función $f(x)=e^x$ en el espacio $S=\text{gn}(1,x,x^2)$ relativa a la norma inducida por el producto interior $\langle g,h\rangle=\int_{-1}^{1}g(x)h(x)dx$ en $C[-1,1]$. Se construye primero una base ortonormal $\{u_1,u_2,u_3\}$ para el espacio S a partir de los generadores L.I. 1, x y x^2:

$$\|1\|^2=\int_{-1}^{1}dx=2$$

y, por tanto,

$$u_1=\frac{1}{\sqrt{2}}.$$

$$\begin{aligned} w_2 &= x-\langle x,u_1\rangle u_1 \\ &= x-\frac{1}{2}\int_{-1}^{1}x \\ &= x, \end{aligned}$$

$$\begin{aligned} \|w_2\|^2 &= \int_{-1}^{1}x^2dx \\ &= \frac{2}{3} \end{aligned}$$

y, entonces,

$$u_2=\sqrt{\frac{3}{2}}x.$$

$$\begin{aligned} w_3 &= x^2-\left\langle x^2,u_1\right\rangle u_1-\left\langle x^2,u_2\right\rangle u_2 \\ &= x^2-\frac{1}{2}\int_{-1}^{1}x^2dx-\frac{3}{2}\int_{-1}^{1}x^3dx \\ &= x^2-\frac{1}{3}, \end{aligned}$$

$$\begin{aligned} \|w_3\|^2 &= \int_{-1}^{1}\left(x^2-\frac{1}{3}\right)^2dx \\ &= \frac{8}{45} \end{aligned}$$

y, por tanto,

$$u_3=\frac{3}{2}\sqrt{\frac{5}{2}}\left(x^2-\frac{1}{3}\right).$$

Del teorema 4.11 (cfr. pág. 269) la aproximación óptima está dada por

$$p=\sum_{i=1}^{3}\langle f,u_i\rangle u_i.$$

$$\begin{aligned} \left\langle e^x,1/\sqrt{2}\right\rangle &= \tfrac{1}{\sqrt{2}}\int_{-1}^{1}e^{-x}dx \\ &= \tfrac{1}{2}\sqrt{2}\left(-e^{-1}+e\right), \end{aligned}$$

$$\left\langle e^x, \sqrt{\tfrac{3}{2}}x \right\rangle = \sqrt{\tfrac{3}{2}} \int_{-1}^{1} xe^x dx$$
$$= \sqrt{6}e^{-1},$$

$$\left\langle e^x, \tfrac{3}{2}\sqrt{\tfrac{5}{2}}\left(x^2 - \tfrac{1}{3}\right) \right\rangle = \tfrac{3}{2}\sqrt{\tfrac{5}{2}} \int_{-1}^{1} e^x \left(x^2 - \tfrac{1}{3}\right) dx$$
$$= \tfrac{3}{4}\sqrt{10}\left(\tfrac{2}{3}e - \tfrac{14}{3}e^{-1}\right).$$

Luego,

$$p = \tfrac{1}{2}\sqrt{2}\left(-e^{-1} + e\right)\tfrac{1}{\sqrt{2}} + \left(\sqrt{6}e^{-1}\right)\sqrt{\tfrac{3}{2}}x$$
$$+ \tfrac{3}{4}\sqrt{10}\left(\tfrac{2}{3}e - \tfrac{14}{3}e^{-1}\right)\left(\tfrac{3}{2}\sqrt{\tfrac{5}{2}}\left(x^2 - \tfrac{1}{3}\right)\right)$$
$$= \tfrac{33}{4}e^{-1} - \tfrac{3}{4}e + 3e^{-1}x + \left(\tfrac{15}{4}e - \tfrac{105}{4}e^{-1}\right)x^2.$$

Entonces,

$$a = \tfrac{33}{4}e^{-1} - \tfrac{3}{4}e,$$
$$b = 3e^{-1},$$
$$c = \tfrac{15}{4}e - \tfrac{105}{4}e^{-1}. \quad ✔$$

Espacios vectoriales normados

23 Calcular $\|f\|_\infty$, si $f(x) = \dfrac{x}{x^2+4}$ en $C[-3,4]$.

Solución El máximo absoluto de $|f|$ se alcanza en los extremos del intervalo $[-3,4]$ o en los puntos críticos de la función f que están en el intervalo $(-3,4)$.

$$f'(x) = \frac{4-x^2}{x^2+4} = 0 \Rightarrow x = \pm 2.$$

$$|f(-3)| = \left|\frac{-3}{(-3)^2+4}\right| = \frac{3}{13},$$
$$|f(-2)| = \left|\frac{2}{(-2)^2+4}\right| = \frac{1}{4} = |f(2)|,$$
$$|f(4)| = \left|\frac{4}{(4)^2+4}\right| = \frac{1}{5}.$$

Por tanto,

$$\|f\|_\infty = \frac{1}{4}. \quad ✔$$

24 Encontrar la distancia entre las funciones $f(x) = 2\operatorname{sen} x$ y $g(x) = -\cos 2x$ relativa a la norma uniforme $\|\cdot\|$ en $C[0,\pi]$.

Solución $$\mathrm{d}(f,g) = \|f - g\|_\infty.$$

$$(f(x)-g(x))' = (2\operatorname{sen} x+\cos 2x)'$$
$$= 2\cos x - 2\operatorname{sen}(2x).$$
$$2\cos x - 2\operatorname{sen}(2x) = 0 \Rightarrow$$
$$2\cos x - 4\cos x \operatorname{sen} x = 0 \Rightarrow$$
$$2\cos(x)(1-2\operatorname{sen} x) = 0.$$

Los puntos críticos de $f-g$ en $(0,\pi)$ son entonces $x=\frac{\pi}{2}$, $x=\frac{\pi}{6}$ y $\frac{5\pi}{6}$. Puesto que,

$$\left|2\operatorname{sen}\frac{\pi}{2}+\cos 2\frac{\pi}{2}\right| = 1,$$
$$\left|2\operatorname{sen}\frac{\pi}{6}+\cos 2\frac{\pi}{6}\right| = \frac{3}{2},$$
$$\left|2\operatorname{sen}\frac{5\pi}{6}+\cos 2\frac{5\pi}{6}\right| = \frac{3}{2},$$
$$\left|2\operatorname{sen}\frac{5\pi}{6}+\cos 2\frac{5\pi}{6}\right| = 1 \quad \text{y}$$
$$\left|2\operatorname{sen}\frac{5\pi}{6}+\cos 2\frac{5\pi}{6}\right| = 1,$$

se tiene

$$\mathrm{d}(f,g) = \frac{3}{2}. \quad ✔$$

En los ejercicios 25 y 26, $\|f\|_1 = \int_a^b |f(x)|\,dx$ en el espacio $C[a,b]$.

25 Calcular $\|f\|_1$ si $f(x) = \dfrac{x}{1+x^2}$, $a=-1$ y $b=1$.

Solución

$$\begin{aligned}\|f\|_1 &= \int_{-1}^{1} |f(x)|\,dx \\ &= \int_{-1}^{0} \frac{-x}{1+x^2}dx + \int_0^1 \frac{x}{1+x^2}dx \\ &= -\frac{1}{2}\ln(1+x^2)\Big|_{-1}^{0} + -\frac{1}{2}\ln(1+x^2)\Big|_{0}^{1} \\ &= \ln(2). \quad ✔\end{aligned}$$

26 Calcular $\mathrm{d}(f,g)$ en $C[a,b]$ relativa a la norma $\|f\|_1 = \displaystyle\int_a^b |f(x)|\,dx$ si $f(x)=\operatorname{sen} x$ y $g(x)=\cos x$ y $a=0$, $b=\pi$.

Solución

$$\begin{aligned}\mathrm{d}(f,g) &= \|f-g\|_1 \\ \int_0^{\pi} |\operatorname{sen} x - \cos x|\,dx &= \int_0^{\pi/4} (\cos x - \operatorname{sen} x)dx + \int_{\pi/4}^{\pi} (\operatorname{sen} x - \cos x)dx \\ &= [\operatorname{sen} x + \cos x]_0^{\pi/4} + [-\cos x - \operatorname{sen} x]_{\pi/4}^{\pi} \\ &= \sqrt{2} - 1 + 1 + \sqrt{2} \\ &= 2\sqrt{2}. \quad ✔\end{aligned}$$

27 Un conjunto no vacío de números reales A es acotado si existe $M \geq 0$ tal que $|x| \leq M$ para todo $x \in A$. Se dice entonces que M es una cota de A. A la menor cota s de un conjunto acotado, se le llama el supremo de este conjunto y se denota por

$$s = \sup A.$$

(a) Sea $B(\mathbb{R})$ el conjunto de todas las sucesiones de números reales que son acotadas. Probar que $B(\mathbb{R})$ es un subespacio de $\mathbb{R}^{\infty}$, el espacio de sucesiones reales.

(b) Si $(a_n) \in B(\mathbb{R})$ se define

$$\|(a_n)\| = \sup\{|a_n| : n \in \mathbb{N}\}.$$

Probar que $\|(a_n)\|$ es una norma en $B(\mathbb{R})$.

DEMOSTRACIÓN ■ *(a)* Obviamente la sucesión constante cero es acotada. Sean $(a_n), (b_n) \in B(\mathbb{R})$ y $M_1, M_2 \geq 0$ cotas de sendas sucesiones, entonces

$$|a_n + b_n| \leq |a_n| + |b_n| \leq M_1 + M_2 \; \forall n$$

Así que la sucesión $(a_n + b_n)$ es también acotada. Si $\alpha \in \mathbb{R}$, entonces

$$|\alpha a_n| = |\alpha|\,|a_n| \leq |\alpha| M_1 \; \forall n$$

y, por tanto, la sucesión (αa_n) es acotada, lo cual prueba que $B(\mathbb{R})$ es un subespacio de $\mathbb{R}^{\infty}$.

(b) Sean $(a_n), (b_n) \in B(\mathbb{R})$ y $\alpha \in \mathbb{R}$.

(i) Puesto que

$$0 \leq |a_n| \leq \|(a_n)\| \; \forall n$$

se tiene $\|(a_n)\| \geq 0$.

(ii) Es claro que la norma de la sucesión constante cero es cero. Si $\|(a_n)\| = 0$, entonces

$$0 \leq |a_n| \leq \|(a_n)\| = 0 \; \forall n$$

de donde se desprende que $a_n = 0$ para todo n y, por tanto, (a_n) es la sucesión constante cero.

(iii) Si $\alpha = 0$, es claro que $\|\alpha(a_n)\| = |\alpha|\,\|(a_n)\| = 0$. Suponga que $\alpha \neq 0$; puesto que $|a_n| \leq \|(a_n)\|$ para todo n, se cumple

$$|\alpha a_n| = |\alpha|\,|a_n| \leq |\alpha|\,\|(a_n)\| \; \forall n,$$

así que $|\alpha|\,\|(a_n)\|$ es una cota de la sucesión (αa_n); y puesto que $\|(\alpha a_n)\|$ es la menor cota,

$$\|(\alpha a_n)\| \leq |\alpha|\,\|(a_n)\|.$$

Por otra parte,

$$|\alpha|\,|a_n| = |\alpha a_n| \leq \|(\alpha a_n)\| \; \forall n \quad \therefore$$

$$|a_n| \leq \frac{1}{|\alpha|}\|(\alpha a_n)\| \; \forall n$$

y, por ende,

$$\|(a_n)\| \leq \frac{1}{|\alpha|}\|(\alpha a_n)\|.$$

Luego,

$$|\alpha|\,\|(a_n)\| \leq \|(\alpha a_n)\| \leq |\alpha|\,\|(a_n)\|$$

lo cual implica

$$\|(\alpha a_n)\| = |\alpha|\,\|(a_n)\|.$$

(iv) Puesto que $|a_n| \leq \|(a_n)\|$ y $|b_n| \leq \|(b_n)\|$ para todo n, se tiene

$$|a_n + b_n| \leq \|(a_n)\| + \|(b_n)\| \ \forall n.$$

Entonces, $\|(a_n)\| + \|(b_n)\|$ es una cota para $(a_n + b_n)$ y, por tanto, ya que $\|(a_n + b_n)\|$ es la menor cota,

$$\|(a_n + b_n)\| \leq \|(a_n)\| + \|(b_n)\|. \quad \blacksquare$$

En los ejercicios 28 a 29 *(i)* probar que la sucesión (a_n) pertenece a $B(\mathbb{R})$ y *(ii)* calcular $\|(a_n)\|$.

28 $a_n = \frac{1}{n} + (-1)^n$.

Solución *(i)* Para todo $n \in \mathbb{N}$,

$$\begin{aligned} \left|\frac{1}{n} + (-1)^n\right| &\leq \frac{1}{n} + |(-1)^n| \\ &= \frac{1}{n} + 1 \\ &\leq 2. \end{aligned}$$

y, por tanto, (a_n) es acotada. *(ii)* Para cada n:

$$\left|\frac{1}{2n} + (-1)^{2n}\right| = 1 + \frac{1}{2n} \leq \frac{3}{2}$$

y

$$\begin{aligned} \left|\frac{1}{2n+1} + (-1)^{2n+1}\right| &= \left|\frac{1}{2n+1} - 1\right| \\ &= 1 - \frac{1}{2n+1} \leq \frac{2}{3} \end{aligned}$$

Puesto que todo número natural es par o impar, se concluye que

$$|a_n| \leq \frac{3}{2}.$$

Y ya que $|a_2| = \frac{3}{2}$, se tiene

$$\|(a_n)\| = \frac{3}{2}. \quad ✔$$

29 $a_n = \dfrac{n}{n+1}$.

Solución *(i)* Para todo $n \in \mathbb{N}$:

$$|a_n| = \frac{n}{n+1} \leq 1.$$

Por tanto, 1 es cota de (a_n). *(ii)* Suponga que M es cota de (a_n) con $0 < M < 1$. Ya que $\text{lím}_{n\to\infty} \frac{n}{n+1} = 1$, existe n_0 tal que

$$1 - \frac{n_0}{n_0+1} = \left|1 - \frac{n_0}{n_0+1}\right| < 1 - M$$

de donde:

$$M < \frac{n_0}{n_0+1},$$

por lo que M no puede existir. Por ende, la menor cota de (a_n) es 1; entonces

$$\|(a_n)\| = 1. \quad ✔$$

30 Sea $\mathfrak{M}_{m\times n}$ el espacio vectorial de las matrices de tamaño $m \times n$ y $A \in \mathfrak{M}_{m\times n}$. Si $\|\cdot\|_2$ representa la norma euclidiana en $\mathbb{R}^k$; esto es, $\|(x_1, \ldots, x_k)\|_2 = \left(\sum_{i=1}^k x_i^2\right)^{1/2}$, para $k = n, m$, probar que el conjunto

$$S_A = \{\|A\vec{x}\|_2 : \vec{x} \in \mathbb{R}^n,\ \|\vec{x}\|_2 = 1\}$$

es acotado.

DEMOSTRACIÓN ■ Sean $\vec{e}_i$, $i = 1, \ldots, n$, la base canónica de $\mathbb{R}^n$ y $M = \text{máx}_{1\leq i\leq n} \|A\vec{e}_i\|_2$. Si $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ y $\|\vec{x}\|_2 = 1$, entonces

$$\begin{aligned}
\|A\vec{x}\|_2 &= \|A(x_1\vec{e}_1 + \cdots + x_n\vec{e}_n\|_2 \\
&\leq |x_1|\,\|A\vec{e}_1\|_2 + \cdots + |x_n|\,\|A\vec{e}_n\|_2 \\
&\leq (|x_1| + \cdots + |x_n|)\,M \\
&\leq Mn\,\|\vec{x}\|_\infty \\
&\leq Mn\,\|\vec{x}\|_2 \\
&= Mn. \quad ■
\end{aligned}$$

31 Sean $A \in \mathfrak{M}_{m\times n}$, $\|\cdot\|_2$ y S_A como en el ejercicio precedente. Se define

$$\|A\| = \sup S_A.$$

(a) Mostrar que $\|A\vec{x}\|_2 \leq \|A\|\,\|\vec{x}\|_2$ para todo $\vec{x} \in \mathbb{R}^n$.

(b) Probar que $\|A\|$ es una norma en $\mathfrak{M}_{m\times n}$.

DEMOSTRACIÓN ■ Sean $A, B \in \mathfrak{M}_{m\times n}$ y $\alpha \in \mathbb{R}$.

(a) Si $\vec{x} = \vec{0}_{\mathbb{R}^n}$, entonces $\|A\vec{x}\|_2 = \left\|\vec{0}_{\mathbb{R}^m}\right\|_2 = 0 = \|A\|\,\|\vec{x}\|$. Sean $\vec{x} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$ y $\vec{x}_1 = \frac{1}{\|\vec{x}\|_2}\vec{x}$, entonces $\|\vec{x}_1\|_2 = 1$ y, por tanto,

$$\begin{aligned}\frac{1}{\|\vec{x}\|_2}\|A\vec{x}\|_2 &= \left\|A\left(\tfrac{1}{\|\vec{x}\|_2}\vec{x}\right)\right\|_2\\ &= \|A\vec{x}_1\|_2\\ &\le \|A\|,\end{aligned}$$

lo cual implica

$$\|A\vec{x}\|_2 \le \|A\|\,\|\vec{x}\|_2$$

(b) *(i)* Si $\vec{x}\in\mathbb{R}^n$ y $\|\vec{x}\|_2 = 1$, entonces

$$0 \le \|A\vec{x}\|_2 \le \|A\|.$$

(ii) Si $A = \mathscr{O}$, la matriz cero, entonces,

$$\|A\vec{x}\|_2 = \left\|\vec{0}_{\mathbb{R}^m}\right\| = 0$$

para todo $\vec{x}\in\mathbb{R}^n$, con $\|\vec{x}\|_2 = 1$ y por tanto $\|A\| = 0$. Suponga que $\|A\| = 0$, entonces

$$\|A\vec{x}\|_2 \le \|A\|\,\|\vec{x}\|_2 = 0\cdot\|\vec{x}\|_2 = 0$$

para todo $\vec{x}\in\mathbb{R}^n$; de donde se desprende que

$$A\vec{x} = \vec{0}_{\mathbb{R}^m}$$

para todo $\vec{x}\in\mathbb{R}^n$. Del lema 5.1 (cfr. pág 436) se concluye que $A = \mathscr{O}$.

(iii) Si $\alpha = 0$, entonces $\|\alpha A\| = \|\mathscr{O}\| = 0 = |\alpha|\,\|A\|$. Si $\alpha\ne 0$, entonces, para cada $\vec{x}\in\mathbb{R}^n$, con $\|\vec{x}\|_2 = 1$, se tiene

$$\begin{aligned}\|(\alpha A)\vec{x}\|_2 &= \|A(\alpha\vec{x})\|_2\\ &\le \|A\|\,\|\alpha\vec{x}\|_2\\ &= \|A\|\,|\alpha|\,\|\vec{x}\|_2\\ &= |\alpha|\,\|A\|\end{aligned}$$

y, por tanto,

$$\|\alpha A\| \le |\alpha|\,\|A\|;$$

y

$$\begin{aligned}|\alpha|\,\|A\vec{x}\|_2 &= \|A(\alpha\vec{x})\|_2\\ &= \|(\alpha A)\vec{x}\|_2\\ &\le \|\alpha A\|;\end{aligned}$$

de donde,

$$\|A\vec{x}\|_2 \le \frac{1}{|\alpha|}\|\alpha A\|.$$

Entonces,

$$\|A\| \le \frac{1}{|\alpha|}\|\alpha A\|;$$

luego,

$$|\alpha|\,\|A\| \leq \|\alpha A\| \leq |\alpha|\,\|A\|\,;$$

así que $\|\alpha A\| = |\alpha|\,\|A\|$.

(iv) Para todo vector unitario $\vec{x} \in \mathbb{R}^n$ ($\|\vec{x}\|_2 = 1$), se tiene

$$\begin{aligned}\|(A+B)\vec{x}\|_2 &= \|A\vec{x}+B\vec{x}\|_2\\ &\leq \|A\vec{x}\|_2 + \|B\vec{x}\|_2\\ &\leq \|A\| + \|B\|\end{aligned}$$

y, por tanto,

$$\|A+B\| \leq \|A\| + \|B\|\,. \quad ■$$

32 Sean $A \in \mathfrak{M}_{m\times n}$, $B \in \mathfrak{M}_{n\times p}$ y $\|C\| = \sup S_C$ la norma del ejercicio precedente. Probar que

$$\|AB\| \leq \|A\|\,\|B\|\,.$$

DEMOSTRACIÓN ■ Sea $\vec{x} \in \mathbb{R}^p$ con $\|\vec{x}\|_2 = 1$, entonces

$$\begin{aligned}\|(AB)\vec{x}\|_2 &= \|A(B\vec{x})\|_2\\ &\leq \|A\|\,\|B\vec{x}\|_2\\ &\leq \|A\|\,\|B\|\,,\end{aligned}$$

por lo que

$$\|AB\| \leq \|A\|\,\|B\|\,. \quad ■$$

En los ejercicios 33 a 35, $C[\alpha,\beta]$ es el espacio de funciones continuas en el intervalo $[\alpha,\beta]$, $p > 1$ es un número real y, para cada $f \in C[\alpha,\beta]$,

$$\|f\|_p = \left(\int_\alpha^\beta |f(x)|^p\,dx\right)^{1/p}$$

(puesto que $f \in C[\alpha,\beta]$, la función $|f|^p$ también y por tanto es integrable en $[\alpha,\beta]$).

33 Si $f,g \in C[\alpha,\beta]$ y $p,p^* > 1$ son índices conjugados ($\frac{1}{p}+\frac{1}{p^*} = 1$), probar que

$$\int_\alpha^\beta |f(x)g(x)|dx \leq \|f\|_p\,\|g\|_{p^*}\,.$$

Es decir,

$$|\langle f,g\rangle| \leq \|f\|_p\,\|g\|_{p^*}$$

donde $\langle f,g\rangle = \int_\alpha^\beta f(x)g(x)dx$ es el producto interior usual en $C[\alpha,\beta]$.

DEMOSTRACIÓN ■ Si f o g es la función constante cero en $[\alpha,\beta]$, la afirmación es evidentemente verdadera. Suponga que ambas funciones no son nulas en $[\alpha,\beta]$; entonces, por continuidad, $\|f\|_p \neq 0 \neq \|g\|_{p^*}$. Sean $x \in [\alpha,\beta]$ un punto fijo de este intervalo y

$$a = \frac{|f(x)|^p}{\|f\|_p^p} \quad \text{y}$$

$$b = \frac{|g(x)|^{p^*}}{\|g\|_{p^*}^{p^*}}$$

en el lema 4.2 (cfr. pág. 329). Entonces, por (4.47) de este lema

$$\frac{|f(x)|}{\|f\|_p}\frac{|g(x)|}{\|g\|_{p^*}} \leq \frac{1}{p}\frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{p^*}\frac{|g(x)|^{p^*}}{\|g\|_{p^*}^{p^*}}$$

para cada $x \in [\alpha, \beta]$. Luego,

$$\begin{aligned}\frac{1}{\|f\|_p \|g\|_{p^*}} \int_\alpha^\beta |f(x)|\,|g(x)|\,dx &\leq \frac{1}{p\,\|f\|_p^p} \int_\alpha^\beta |f(x)|^p\,dx \\ &\quad + \frac{1}{p^*\,\|f\|_{p^*}^{p^*}} \int_\alpha^\beta |g(x)|^{p^*}\,dx \\ &= \frac{1}{p\,\|f\|_p^p}\,\|f\|_p^p + \frac{1}{p^*\,\|g\|_{p^*}^{p^*}}\,\|g\|_{p^*}^{p^*} \\ &= \frac{1}{p} + \frac{1}{p^*} \\ &= 1\end{aligned}$$

y, por tanto,

$$\int_\alpha^\beta |f(x)|\,|g(x)|\,dx \leq \|f\|_p\,\|g\|_{p^*}\,.$$

Finalmente,

$$\begin{aligned}|\langle f, g\rangle| &= \left|\int_\alpha^\beta f(x)g(x)dx\right| \\ &\leq \int_\alpha^\beta |f(x)|\,|g(x)|\,dx \\ &\leq \|f\|_p\,\|g\|_{p^*}\,. \quad \blacksquare\end{aligned}$$

34 Si $f, g \in C[\alpha, \beta]$, demostrar que

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p\,.$$

DEMOSTRACIÓN ■ Para cualquier $x \in C[\alpha, \beta]$

$$\begin{aligned}|f(x) + g(x)|^p &= |f(x) + g(x)|\,|f(x) + g(x)|^{p-1} \\ &\leq |f(x)|\,|f(x) + g(x)|^{p-1} + |g(x)|\,|f(x) + g(x)|^{p-1}\end{aligned}$$

y, por tanto,

$$\begin{aligned}\int_\alpha^\beta |f(x) + g(x)|^p\,dx &\leq \int_\alpha^\beta |f(x)|\,|f(x) + g(x)|^{p-1}\,dx \\ &\quad + \int_\alpha^\beta |g(x)|\,|f(x) + g(x)|^{p-1}\,dx.\end{aligned}$$

Por el ejercicio precedente,

$$\int_\alpha^\beta |f(x)|\,|f(x)+g(x)|^{p-1}\,dx \le \|f\|_p \left\|(f+g)^{p-1}\right\|_{p^*} \quad \text{y}$$
$$\int_\alpha^\beta |g(x)|\,|f(x)+g(x)|^{p-1}\,dx \le \|g\|_p \left\|(f+g)^{p-1}\right\|_{p^*}.$$

Entonces,

$$\begin{aligned}\int_\alpha^\beta |f(x)+g(x)|^p\,dx &\le \|f\|_p \left\|(f+g)^{p-1}\right\|_{p^*} + \|g\|_p \left\|(f+g)^{p-1}\right\|_{p^*}\\ &= \left(\|f\|_p+\|g\|_p\right)\left\|(f+g)^{p-1}\right\|_{p^*}\\ &= \left(\|f\|_p+\|g\|_p\right)\left(\int_\alpha^\beta |f(x)+g(x)|^{(p-1)p^*}\,dx\right)^{1/p^*}\\ &= \left(\|f\|_p+\|g\|_p\right)\left(\int_\alpha^\beta |f(x)+g(x)|^{p}\,dx\right)^{1/p^*}.\end{aligned}$$

Así que,

$$\left(\int_\alpha^\beta |f(x)+g(x)|^p\,dx\right)^{1-1/p^*} \le \|f\|_p+\|g\|_p$$

esto es,

$$\left(\int_\alpha^\beta |f(x)+g(x)|^p\,dx\right)^{1/p} \le \|f\|_p+\|g\|_p.$$

O sea

$$\|f+g\|_p \le \|f\|_p+\|g\|_p. \quad \blacksquare$$

35 Demostrar que

$$\|f\|_p = \left(\int_\alpha^\beta |f(x)|^p\,dx\right)^{1/p}$$

es una norma en $C[\alpha,\beta]$.

DEMOSTRACIÓN ■ Sean $f,g\in C[\alpha,\beta]$, $\lambda\in\mathbb{R}$.

(a) Es evidente que $\|f\|_p \ge 0$.

(b) Si $f=\theta$, la función constante cero en $[\alpha,\beta]$, es claro que $\|f\|_p=0$. Suponga que $\|f\|_p=0$, entonces

$$\int_\alpha^\beta |f(x)|^p = 0,$$

por continuidad[34] se desprende que $|f(x)|^p=0$ y, por tanto, $f(x)=0$ para todo $x\in[\alpha,\beta]$.

[34] A lo largo de este texto, la demostración de que una función no negativa, continua y cuya integral en un intervalo es cero tiene que ser una función nula se ha realizado varias veces y por esta razón ya no se repite aquí.

(c)
$$
\begin{aligned}
\|\lambda f\|_p &= \left(\int_\alpha^\beta |\lambda f(x)|^p\,dx\right)^{1/p}\\
&= \left(\int_\alpha^\beta |\lambda|^p\,|f(x)|^p\,dx\right)^{1/p}\\
&= \left(|\lambda|^p\int_\alpha^\beta |f(x)|^p\,dx\right)^{1/p}\\
&= |\lambda|\left(\int_\alpha^\beta |f(x)|^p\,dx\right)^{1/p}\\
&= |\lambda|\,\|f\|_p\,.
\end{aligned}
$$

(d) Por el ejercicio anterior

$$\|f+g\|_p \le \|f\|_p + \|g\|_p\,. \quad \blacksquare$$

36 Calcular, en relación con el ejercicio precedente, $\|f\|_p$ si $f(x)=\cos x$, $p=3$, $\alpha=0$ y $\beta=\pi$.

Solución
$$
\begin{aligned}
\|f\|_3^3 &= \int_0^\pi |\cos x|^3\,dx\\
&= \int_0^{\pi/2} \cos^3 x\,dx - \int_{\pi/2}^{\pi} \cos^3 x\,dx.
\end{aligned}
$$

$$
\begin{aligned}
\int \cos^3 x\,dx &= \int \cos^2 x \cos x\,dx\\
&= \int (1-\operatorname{sen}^2 x)\cos x\,dx\\
&= \int \cos x\,dx - \int \operatorname{sen}^2 x\cos x\,dx\\
&= \operatorname{sen} x - \frac{\operatorname{sen}^3 x}{3} + C
\end{aligned}
$$

Por tanto,

$$
\begin{aligned}
\|f\|_3^3 &= \left[\operatorname{sen} x - \frac{\operatorname{sen}^3 x}{3}\right]_0^{\pi/2} - \left[\operatorname{sen} x - \frac{\operatorname{sen}^3 x}{3}\right]_{\pi/2}^{\pi}\\
&= \frac{2}{3} + \frac{2}{3}\\
&= \frac{4}{3}
\end{aligned}
$$

y, entonces,

$$\|f\|_3 = \sqrt[3]{\frac{4}{3}}. \quad ✔$$

37 Sean $a\in\mathbb{R}$ y $\mathscr{L}_1[a,\infty)$ el conjunto de funciones continuas $f:[a,\infty)\to\mathbb{R}$ tales que $\int_a^\infty |f(x)|\,dx$ converge.

(a) Probar que $\mathscr{L}_1[a,\infty)$ es un espacio vectorial con las operaciones suma de funciones y multiplicación de un escalar con una función usuales.

(b) Sea $\|f\|_1 = \int_a^\infty |f(x)|\,dx$, para cada $f\in\mathscr{L}_1[a,\infty)$. Mostrar que $\|f\|_1$ es una norma en $\mathscr{L}_1[a,\infty)$.

DEMOSTRACIÓN ■ *(a)* Dado que $\mathscr{L}_1[a,\infty) \subset C[a,\infty)$, basta probar que $\mathscr{L}_1[a,\infty) < C[a,\infty)$.

(i) Es evidente que la función constante cero pertenece a $\mathscr{L}_1[a,\infty)$.

(ii) Sean $f, g \in \mathscr{L}_1[a,\infty)$, entonces, para todo $x \in [a,\infty)$,

$$|f(x)+g(x)| \leq |f(x)|+|g(x)|$$

y, por tanto, para todo $r \geq a$,

$$\begin{aligned} \int_a^r |f(x)+g(x)|\,dx &\leq \int_a^r |f(x)|\,dx + \int_a^r |f(x)|\,dx \\ &\leq \int_a^\infty |f(x)|\,dx + \int_a^\infty |f(x)|\,dx, \end{aligned}$$

de donde $\int_a^\infty |f(x)+g(x)|\,dx$ converge, entonces $f+g \in \mathscr{L}_1[a,\infty)$.

(iii) Sea $\alpha \in \mathbb{R}$ y $f \in \mathscr{L}_1[a,\infty)$, entonces

$$\begin{aligned} \lim_{r\to\infty} \int_a^r |\alpha f(x)|\,dx &= |\alpha| \lim_{r\to\infty} \int_a^r |f(x)|\,dx \\ &= |\alpha| \int_a^\infty |f(x)|\,dx \end{aligned}$$

y, por tanto, $\int_a^\infty |\alpha f(x)|\,dx$ converge, así que $\alpha f \in \mathscr{L}_1[a,\infty)$ y $\|\alpha f\|_1 = |\alpha|\,\|f\|_1$.

(b) Sean $f, g \in \mathscr{L}_1[a,\infty)$ y $\alpha \in \mathbb{R}$.

(i)
$$\|f\|_1 = \int_a^\infty |f(x)|\,dx \geq 0.$$

(ii) Es claro que $\|\theta\|_1 = 0$ para la función constante cero. Si $\|f\| = 0$, entonces

$$\int_a^r |f(x)|\,dx \leq \int_a^\infty |f(x)|\,dx = 0$$

y, por tanto, $\int_a^r |f(x)|\,dx = 0$; de aquí que, por continuidad, $f(x) = 0$ para todo intervalo $[a,r]$; es decir, f es la función constante cero en $[a,\infty)$.

(iii) Se probó en 37*(a)(iii)*.

(iv) Puesto que

$$\int_a^r |f(x)+g(x)|\,dx \leq \int_a^r |f(x)|\,dx + \int_a^r |g(x)|\,dx$$

se tiene

$$\int_a^\infty |f(x)+g(x)|\,dx \leq \int_a^\infty |f(x)|\,dx + \int_a^\infty |g(x)|\,dx$$

es decir,

$$\|f+g\|_1 \leq \|f\|_1 + \|g\|_1 . \quad ■$$

38 Sean $\mathscr{L}[0,\infty)$ y $\|f\|_1 = \int_0^\infty |f(x)|\,dx$ el espacio y la norma del ejercicio precedente para el caso particular $a=0$. Probar que si $f(x) = e^{-x}$, entonces $f \in \mathscr{L}[0,\infty)$ y calcular $\|f\|_1$.

DEMOSTRACIÓN

$$\begin{aligned}\int_0^{\infty}\left|e^{-x}\right|dx &= \lim_{r\to\infty}\int_0^r e^{-x}dx\\ &= \lim_{r\to\infty}\left[-e^{-x}\right]_0^r\\ &= \lim_{r\to\infty}\left[1-e^{-r}\right]\\ &= 1,\end{aligned}$$

lo cual demuestra que $f \in \mathscr{L}[0,\infty)$ y $\|f\|_1 = 1$. ■

En los ejercicios 39 a 41, $S = \text{gn}((-1,1))$ y $\vec{u} = (0,1)$ en $\mathbb{R}^2$.

39 Hallar la aproximación óptima, $\vec{p}^*$, de $\vec{u}$ en el subespacio S respecto a la norma euclidiana $\|(x,y)\| = \sqrt{x^2+y^2}$.

Solución Los elementos de S tienen la forma $(-a,a)$ con $a \in \mathbb{R}$.

$$\begin{aligned}\|\vec{u}-(-a,a)\|^2 &= \|(a,1-a)\|^2\\ &= a^2+(1-a)^2.\end{aligned}$$

Si $\psi(a) = a^2+(1-a)^2$, $a \in \mathbb{R}$, entonces

$$\begin{aligned}\psi'(a) = 4a-2 = 0 \Rightarrow\\ a = \frac{1}{2}\end{aligned}$$

y, puesto que,

$$\psi''(a) = 4 > 0$$

la función ψ alcanza un mínimo en $a = 1/2$. Luego $\|\vec{u}-(-a,a)\|^2$ es mínima (y, por tanto, $\|\vec{u}-(-a,a)\|$ es mínima) cuando $a = 1/2$. Así la aproximación óptima de $\vec{u}$ en S es

$$\vec{p}^* = (-1/2, 1/2)$$

y

$$\|\vec{u}-\vec{p}^*\| = \sqrt{2}/2.$$

Note que en este caso $\vec{p}^*$ también se puede calcular utilizando la proyección de $\vec{u}$ sobre S, pues la norma euclidiana proviene del producto punto de vectores. ✔

40 Encontrar todas las aproximaciones óptimas de $\vec{u}$ en el subespacio S respecto a la norma cúbica

$$\|(x,y)\|_{\infty} = \text{máx}\{|x|,|y|\}.$$

Solución Los elementos de S tienen la forma $(-a,a)$ con $a \in \mathbb{R}$.

$$\begin{aligned}\|\vec{u}-(-a,a)\|_{\infty} &= \|(a,1-a)\|_{\infty}\\ &= \text{máx}\{|a|,|1-a|\}.\end{aligned}$$

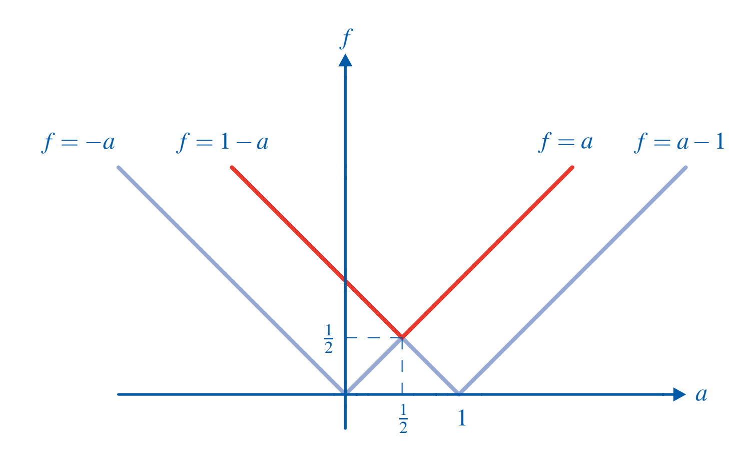


Figura 4-36 •

De la figura 4-36 se desprende que

$$\text{máx}\,\{|a|\,,|1-a|\} = \begin{cases} a, \text{ si } a \geq 1/2 \\ 1-a, \text{ si } a < 1/2 \end{cases}$$

y de la misma gráfica se deduce que

$$\min_{a\in\mathbb{R}} \|\vec{u}-(-a,a)\|_\infty = \min_{a\in\mathbb{R}} \text{máx}\,\{|a|\,,|1-a|\}$$

se alcanza en $a = 1/2$. Así,

$$\vec{p}^{\,*} = (-1/2, 1/2)$$

es la única aproximación óptima de $\vec{u}$ en S y

$$\|\vec{u}-(-1/2,1/2)\|_\infty = 1/2. \quad ✔$$

41 Encontrar todas las aproximaciones óptimas de $\vec{u}$ en el subespacio S respecto a la norma

$$\|(x,y)\|_1 = |x| + |y|.$$

Solución Los elementos de S tienen la forma $(-a,a)$ con $a \in \mathbb{R}$.

$$\begin{aligned} \|\vec{u}-(-a,a)\|_1 &= \|(a,1-a)\|_1 \\ &= |a| + |1-a|. \end{aligned}$$

De la figura 4-36 se desprende que

$$|a| + |1-a| = \begin{cases} 2a-1, & \text{si } a \geq 1 \\ 1, & \text{si } 0 \leq a < 1 \\ 1-2a, & \text{si } a < 0 \end{cases}$$

es decir, para cada $a \in \mathbb{R}$ se tiene

$$\|\vec{u}-(-a,a)\|_1 = \begin{cases} 2a-1, & \text{si } a \geq 1 \\ 1, & \text{si } 0 \leq a < 1 \\ 1-2a, & \text{si } a < 0 \end{cases}$$

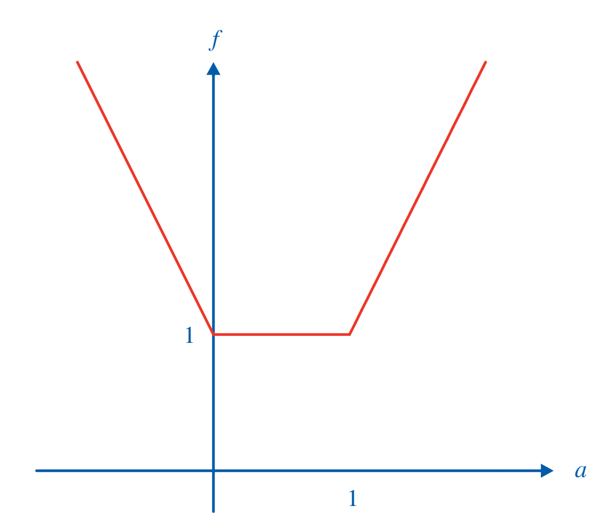


Figura 4-37 •

La figura 4-37 contiene la gráfica de $f = \|\vec{u} - (-a,a)\|_1$ cuando a varía en $\mathbb{R}$; de ella se deduce que

$$\min_{a\in\mathbb{R}} \|\vec{u} - (-a,a)\|_1$$

se alcanza en todos los valores a del intervalo $[0,1]$. Por tanto, $\vec{u}$ tiene una infinidad de aproximaciones óptimas en S; a saber, todos los vectores de la forma

$$\vec{p}^* = a(-1,1) \text{ con } a \in [0,1]$$

y en todos ellos

$$\|\vec{u} - (-a,a)\|_1 = 1. \quad ✔$$

42 Encontrar las aproximaciones óptimas del vector $\vec{u} = (1,0,0)$ en el plano $x - y = 0$, relativas a la norma $\|(x,y,z)\|_1 = |x| + |y| + |z|$ en $\mathbb{R}^3$.

Solución **Un primer método** Este plano es el subespacio $S = \text{gn}((1,1,0),(0,0,1))$. Si $\vec{v} = r(0,0,1) + s(1,1,0) = (s,s,r) \in S$, entonces

$$\begin{aligned}\|\vec{u} - \vec{v}\|_1 &= \|(1,0,0) - (s,s,r)\|_1 \\ &= |s-1| + |s| + |r|.\end{aligned}$$

(a) Suponga $r \geq 0$:

(i) si $s \geq 1$, entonces $|s-1| + |s| + |r| = s - 1 + s + r = 2s - 1 + r$;
(ii) si $0 \leq s < 1$, entonces $|s-1| + |s| + |r| = 1 - s + s + r = 1 + r$;
(iii) si $s < 0$, entonces $|s-1| + |s| + |r| = 1 - s - s + r = 1 - 2s + r$.

(b) Suponga $r < 0$:

(i) si $s \geq 1$, entonces $|s-1| + |s| + |r| = s - 1 + s - r = 2s - 1 - r$;
(ii) si $0 \leq s < 1$, entonces $|s-1| + |s| + |r| = 1 - s + s - r = 1 - r$;
(iii) si $s < 0$, entonces $|s-1| + |s| + |r| = 1 - s - s - r = 1 - 2s - r$.

De esta manera, si

$$\psi(r,s) = \|\vec{u}-\vec{v}\|_1\,,$$

entonces,

$$\frac{\partial\psi}{\partial r}(r,s) = \begin{cases} 1, & \text{si } r>0 \\ -1, & \text{si } r<0 \\ 0, & \text{si } r=0 \end{cases}$$

y

$$\frac{\partial\psi}{\partial s}(r,s) = \begin{cases} 2, & \text{si } s>1 \\ 0, & \text{si } 0\le s\le 1 \\ -2, & \text{si } s<0 \end{cases}$$

Por lo que los puntos críticos de ψ tienen la forma $(s,s,0)$ con $0\le s\le 1$. En todos ellos

$$\|\vec{u}-\vec{v}\|_1 = |s-1|+|s|+|r| = 1$$

y si $r\neq 0$ y/o $s\notin[0,1]$, entonces $\|\vec{u}-\vec{v}\|_1 = |s-1|+|s|+|r|>1$; por lo que las aproximaciones óptimas de $\vec{u}$ en S están dadas por los vectores

$$\vec{p}^{\,*} = s(1,1,0) \text{ con } 0\le s\le 1.$$

Un segundo método Este plano es el subespacio $S=\text{gn}((1,1,0),(0,0,1))$. Si $\vec{v}=r(0,0,1)+s(1,1,0)=(s,s,r)\in S$, entonces

$$\begin{aligned}\|\vec{u}-\vec{v}\|_1 &= \|(1,0,0)-(s,s,r)\|_1 \\ &= |s-1|+|s|+|r|\,.\end{aligned}$$

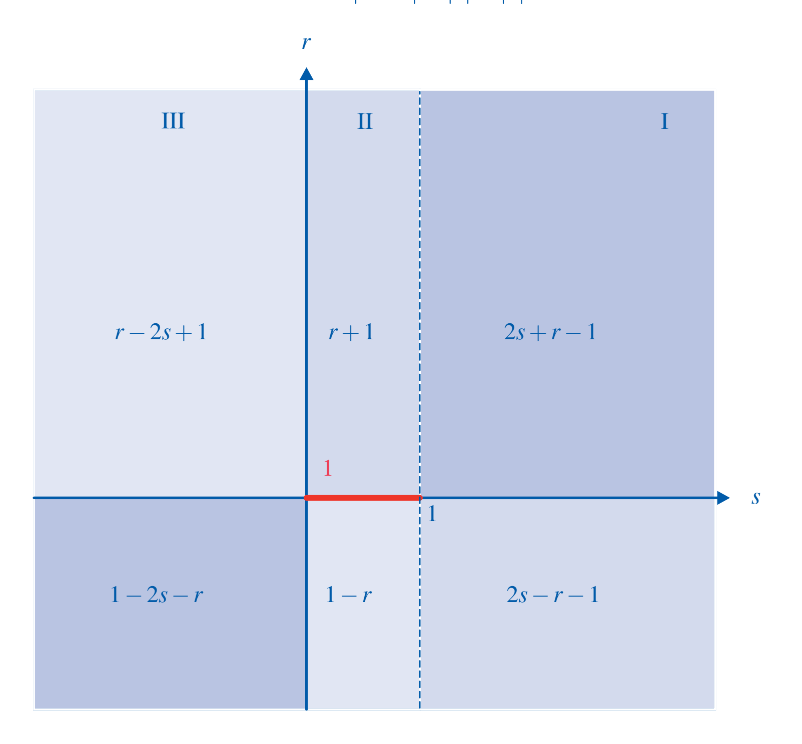


Figura 4-38 •

En la figura 4-38 se ha dividido el plano en zonas sombreadas para evaluar $\|\vec{u}-\vec{v}\|_1 = |s-1|+|s|+|r|$. Por ejemplo, en la zona $s \geq 1$ y $r > 0$, $|s-1|+|s|+|r| = 2s+r-1$; en la zona $0 \leq s < 1$ y $r < 0$, $|s-1|+|s|+|r| = r+1$; etc. En realidad, la información dada en esta figura se puede resumir aún más:

(a) En la franja vertical I, $s \geq 1$,

$$\|\vec{u}-\vec{v}\|_1 = |s-1|+|s|+|r| = 2s-1+|r|.$$

(b) En la franja vertical II, $0 \leq s \leq 1$,

$$\|\vec{u}-\vec{v}\|_1 = |s-1|+|s|+|r| = 1+|r|.$$

(c) En la franja vertical III, $s \leq 0$,

$$\|\vec{u}-\vec{v}\|_1 = |s-1|+|s|+|r| = 1+2|s|+|r|.$$

En la franja I, dado que $s \geq 1$,

$$1 \leq 2s-1+|r|$$

para todo (s,r) en esta franja. Y, puesto que $2s-1+|r| = 1$ cuando $s=1$ y $r=0$, se tiene

$$\text{mín}\{2s-1+|r| : (s,r) \text{ está en la franja I}\} = 1$$

y se alcanza en $(s,r) = (1,0)$.

En la franja III,

$$\|\vec{u}-\vec{v}\|_1 = |s-1|+|s|+|r| = 1+2|s|+|r| \geq 1$$

para todo (s,r) en esta franja y ya que este valor se alcanza en $(0,0)$, se tiene que

$$\text{mín}\{1+2|s|+|r| : (s,r) \text{ está en la franja I}\} = 1$$

y se alcanza en $(0,0)$.

En la franja II, puesto que $0 \leq s \leq 1$,

$$\|\vec{u}-\vec{v}\|_1 = |s-1|+|s|+|r| = 1+|r| \geq 1$$

para todo (s,r) en esta franja y ya que este valor se alcanza en todos los puntos de forma $(s,0)$ con $0 \leq s \leq 1$, se tiene que

$$\text{mín}\{1+|r| : (s,r) \text{ está en la franja I}\} = 1$$

y este valor se alcanza en los puntos $(s,0)$ con $0 \leq s \leq 1$. Así, el conjunto de aproximaciones óptimas para el vector $\vec{u}$ en el espacio S están dados por los vectores

$$\vec{p}^* = s(1,1,0) \text{ con } 0 \leq s \leq 1.$$

y para todos estos puntos

$$\left\|\vec{u}-\vec{p}^*\right\|_1 = 1. \quad ✔$$

43 Encontrar las aproximaciones óptimas del vector $\vec{u} = (1,0,0)$ en el plano $x - y = 0$, relativas a la norma $\|(x,y,z)\|_\infty = \text{máx}\{|x|,|y|,|z|\}$ en $\mathbb{R}^3$.

Solución Este plano es el subespacio $S = \text{gn}((1,1,0),(0,0,1))$. Si $\vec{v} = r(0,0,1) + s(1,1,0) = (s,s,r) \in S$, entonces

$$\begin{aligned}\|\vec{u} - \vec{v}\|_\infty &= \|(1,0,0) - (s,s,r)\|_\infty \\ &= \text{máx}\{|s-1|,|s|,|r|\}.\end{aligned}$$

En la figura 4-39 se ha dividido el plano en zonas sombreadas para calcular

$$\|\vec{u} - \vec{v}\|_\infty = \text{máx}\{|s-1|,|s|,|r|\}.$$

(a) En la franja I, puesto que $s \ge 1$, $|s-1| = s-1$ y $s-1 < s$; por tanto,

$$\|(1,0,0) - (s,s,r)\|_\infty = \text{máx}\{s,|r|\}.$$

(b) En la franja II, ya que $1/2 \le s \le 1$, $|s-1| = 1-s < s$; por tanto,

$$\|(1,0,0) - (s,s,r)\|_\infty = \text{máx}\{s,|r|\}.$$

(c) En la franja III, dado que $0 \le s \le 1/2$, $|s-1| = 1-s$ y $1-s > s$; entonces

$$\|(1,0,0) - (s,s,r)\|_\infty = \text{máx}\{1-s,|r|\}.$$

(d) En la franja IV, ya que $s \le 0$, $|s-1| = 1-s = 1+|s| \ge |s|$ y, por ende,

$$\|(1,0,0) - (s,s,r)\|_\infty = \text{máx}\{1+|s|,|r|\}.$$

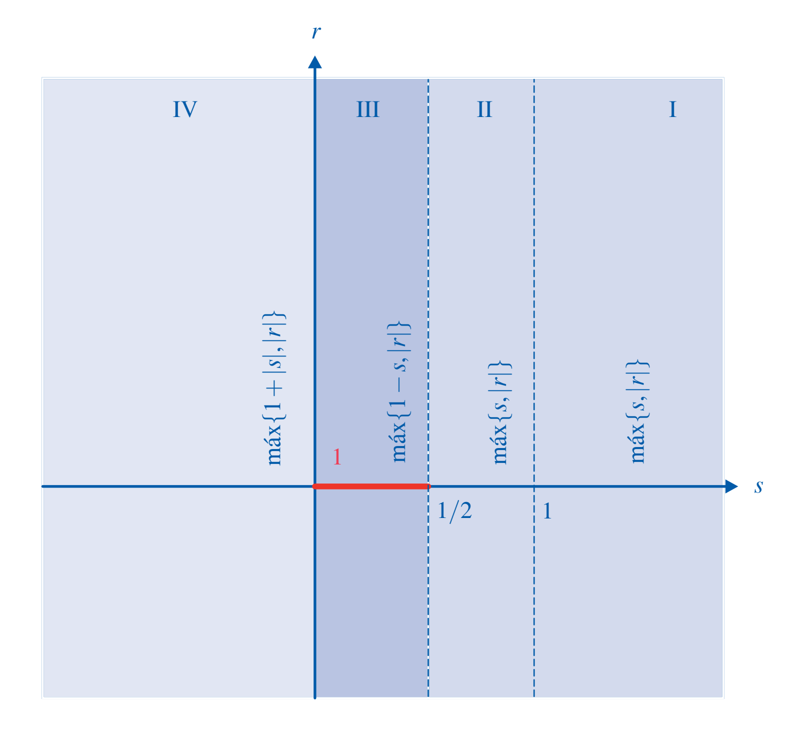


Figura 4-39 •

En todos los puntos (s,r) de las franjas I y II, se tiene

$$1/2 \leq \text{máx}\{s, |r|\}$$

pues $s \geq 1/2$; y ya que el punto $(1/2,0)$ pertenece a la reunión de ambas franjas

$$\text{mín}\,\|(1,0,0)-(s,s,r)\|_\infty = 1/2$$

y se alcanza el punto $(1/2,0)$. En los puntos de la franja III, $0 \leq s \leq 1/2$ y, por tanto, $1/2 \leq 1-s \leq 1$; luego,

$$1/2 \leq \text{máx}\{1-s, |r|\}$$

para todos los puntos en esta franja y puesto que $1/2$ se alcanza en el punto $(s,r)=(1/2,0)$, se concluye que

$$\text{mín}\,\|(1,0,0)-(s,s,r)\|_\infty = 1/2.$$

en esta franja y dicho valor se alcanza en $s=1/2$, $r=0$. En la franja IV

$$1 \leq \text{máx}\{1+|s|, |r|\}$$

para todo par (s,r) en esta franja; y puesto que 1 se alcanza en $(0,0)$, se deduce que

$$\text{mín}\,\|(1,0,0)-(s,s,r)\|_\infty = 1$$

en esta franja y se alcanza en $(s,r)=(0,0)$. De lo precedente se deduce que

$$\|(1,0,0)-(s,s,r)\|_\infty$$

es mínima cuando $s=1/2$ y $r=0$; luego la aproximación óptima para $\vec{u}$ en el subespacio S está dada por

$$\vec{p^*} = (1/2,1/2,0). \quad ✔$$

44 Encontrar una aproximación óptima para la función $f(x)=x$ en el subespacio $S=\text{gn}(1)$ de $C[0,1]$ respecto a la norma

$$\|f\|_4 = \left(\int_0^1 |f(x)|^4\,dx\right)^{1/4}$$

del ejercicio resuelto 35 (con $p=4$).

Solución Sea $a \in \text{gn}(1)$, entonces

$$\begin{aligned}\|x-a\|_4^4 &= \int_0^1 (x-a)^4\,dx \\ &= a^4-2a^3+2a^2-a+\frac{1}{5}.\end{aligned}$$

Si

$$\varphi(a) = a^4-2a^3+2a^2-a+\frac{1}{5},\ a \in \mathbb{R},$$

entonces

$$\begin{aligned}\varphi'(a) &= 4a^3 - 6a^2 + 4a - 1 \\ &= (2a-1)\left(2a^2 - 2a + 1\right)\end{aligned}$$

y

$$\varphi''(1/2) = 1 > 0.$$

Por tanto, φ es mínima en $a = 1/2$; luego una aproximación óptima para $f(x) = x$ en $S = \text{gn}(1)$ es $p^*(x) = 1/2$ y

$$\|x - 1/2\|_4 = \left(\int_0^1 (x - 1/2)^4\, dx\right)^{1/4} = \left(\frac{1}{80}\right)^{1/4}$$

es el valor mínimo que alcanza $\|x - a\|_4$ cuando a varía en $\mathbb{R}$. De lo precedente se desprende que f tiene sólo una aproximación óptima en S. ✔

45 Encontrar la aproximación óptima de $f(x) = e^x$ en el subespacio $S = \text{gn}(1)$ de $C[0,1]$ respecto a la norma uniforme $\|g\|_\infty$.

Solución De la figura 4-40*(i)* se desprende que

$$\|e^x - a\|_\infty = \begin{cases} e - a, \text{ si } a \le (e+1)/2 \\ a - 1, \text{ si } a > (e+1)/2 \end{cases}$$

y de la misma figura *(ii)* se deduce que

$$\min_{a\in\mathbb{R}} \|e^x - a\|_\infty = \min_{a\in\mathbb{R}} \max |e^x - a|$$

se alcanza en $a = (1+e)/2$ y tiene un valor de $(e-1)/2$. Entonces, la aproximación óptima para $f(x) = e^x$ en el subespacio S está dada por

$$p^* = \frac{1+e}{2}$$

y

$$\|f - p^*\|_\infty = \frac{e-1}{2}. \quad ✔$$

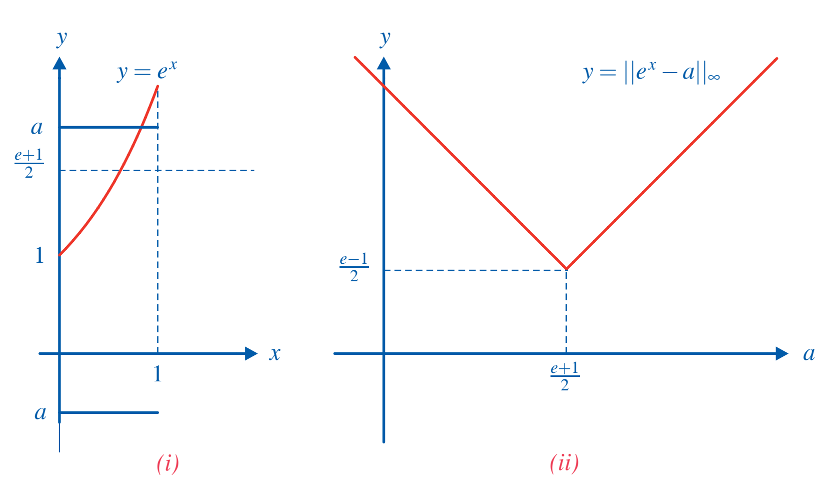


Figura 4-40 •

46 Encontrar las aproximaciones óptimas de $f(x)=e^x$ en el subespacio $S=\text{gn}(1)$ de $C[0,1]$ respecto de la norma $\|g\|_1=\int_0^1|g(x)|\,dx$.

Solución

$$\|f-a\|_1=\int_0^1|e^x-a|\,dx=\begin{cases}\displaystyle\int_0^1(e^x-a)dx, & \text{si } a\le 1\\ \displaystyle\int_0^{\ln a}(a-e^x)dx+\int_{\ln a}^1(e^x-a)dx, & \text{si } 1<a\le e\\ \displaystyle\int_0^1(a-e^x)dx, & \text{si } a>e\end{cases}$$

esto es,

$$\|f-a\|_1=\begin{cases}e-a-1, & \text{si } a\le 1\\ 2a\ln a-3a+1+e, & \text{si } 1<a\le e\\ a-e+1, & \text{si } a>e\end{cases}$$

Así, la función continua $a\mapsto\|f-a\|_1$, $a\in\mathbb{R}$, es decreciente en el intervalo $(-\infty,1)$ (pues es una línea recta con pendiente negativa en éste) y creciente en el intervalo (e,∞) (en éste es una línea recta con pendiente positiva); entonces, los puntos donde alcanza el mínimo son $a=1$, $a=e$ o los puntos críticos de la función $\psi(a)=2a\ln a-3a+1+e$.

$$\begin{aligned}\psi'(a)&=2\ln a-1=0\Rightarrow\\ a&=e^{1/2}\in(1,e);\end{aligned}$$

y ya que

$$\begin{aligned}\|f-1\|_1&=e-2,\\ \|f-e\|_1&=1,\\ \psi(e^{1/2})&=-2e^{1/2}+1+e\end{aligned}$$

se tiene

$$\min_{a\in\mathbb{R}}\|f-a\|_1=1+e-2e^{1/2}$$

y se alcanza en $a=e^{1/2}$. Luego la aproximación óptima de f en S está dada por la función constante $a=e^{1/2}$. ✔

47 Sean $C[a,b]$ el espacio de funciones continuas en el intervalo $[a,b]$, $\|f\|_\infty$ la norma uniforme y $\|f\|=\left(\int_a^b f^2(x)\,dx\right)^{1/2}$ la norma cuadrado medio; demostrar que la convergencia respecto a la norma cuadrado medio no implica la convergencia respecto a la norma uniforme; esto es, si

$$\lim_{n\to\infty}\|f_n-f\|=0$$

entonces, no necesariamente

$$\lim_{n\to\infty}\|f_n-f\|_\infty=0$$

en el espacio $C[a,b]$.

Solución Sean n un entero positivo y f_n la función continua en $[a,b]$, cuya gráfica está contenida en la figura 4-41; esto es

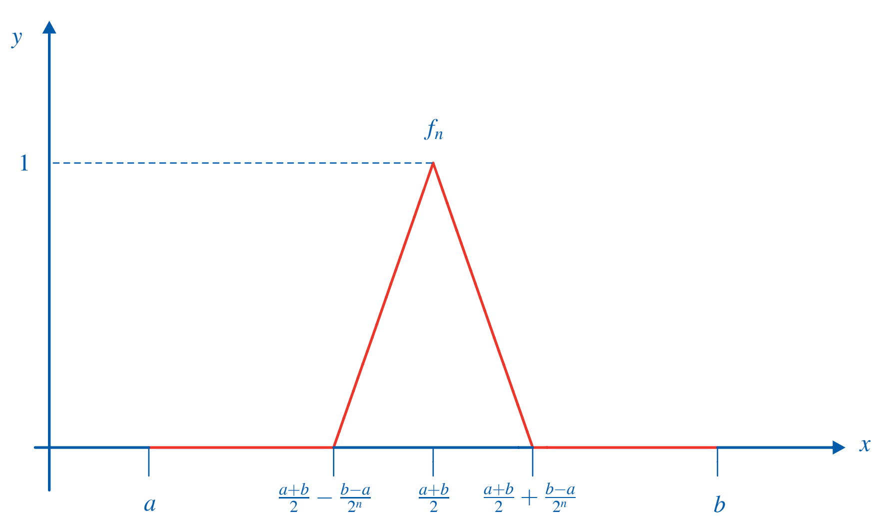


Figura 4-41 •

$$f_n(x)=\begin{cases}\dfrac{2^n}{b-a}x+\left(1-\dfrac{a+b}{b-a}2^{n-1}\right) & \text{si } x\in\left[\dfrac{a+b}{2}-\dfrac{b-a}{2^n},\dfrac{a+b}{2}\right]\\ \dfrac{2^n}{a-b}x+\left(1-\dfrac{a+b}{a-b}2^{n-1}\right) & \text{si } x\in\left[\dfrac{a+b}{2},\dfrac{a+b}{2}+\dfrac{b-a}{2^n}\right]\\ 0 & \text{si } x\in[a,b]-\left[\dfrac{a+b}{2}-\dfrac{b-a}{2^n},\dfrac{a+b}{2}+\dfrac{b-a}{2^n}\right]\end{cases}$$

Es fácil ver que

$$\int_a^b f_n^2(x)dx=\frac{1}{3}\frac{(b-a)}{2^{n-1}}$$

para todo $n=1,2,\ldots$ Por tanto,

$$\lim_{n\to\infty}\|f_n-\theta\|=0,$$

donde la función θ es la función constante cero en el intervalo $[a,b]$. Sin embargo, es claro que

$$\|f_n\|_\infty=1$$

para todo $n=1,2,\ldots$ y, por tanto,

$$\lim_{n\to\infty}\|f_n-\theta\|_\infty=1.$$

Es decir, convergencia en promedio cuadrático no implica convergencia uniforme. ✔

4.3.2 Ejercicios propuestos

El lector encontrará la respuesta a los ejercicios en cursiva en el apéndice E al final del libro.

Espacios con producto interior (respuestas en páginas 1080-1082)

En los ejercicios 1 a 4, determinar si $\langle(x_1,x_2),(y_1,y_2)\rangle$, definido por la fórmula dada, es o no un producto interior en $\mathbb{R}^2$. En caso positivo demostrar rigurosamente que se cumplen las 5 condiciones de la

definición 4.1 (cfr. pág. 236) y en caso negativo mostrar, mediante contraejemplos, las propiedades que no se satisfacen.

1 $\langle (x_1,x_2),(y_1,y_2)\rangle = x_1y_1 - x_2y_2$.

2 $\langle (x_1,x_2),(y_1,y_2)\rangle = x_1y_2 - x_2y_1$.

3 $\langle (x_1,x_2),(y_1,y_2)\rangle = x_1x_2 - y_1y_2$.

4 $\langle (x_1,x_2),(y_1,y_2)\rangle = 2y_1x_1 + y_1x_2 + y_2x_1 + y_2x_2$.

En los ejercicios 5 a 19 probar que la asignación $(\vec{x},\vec{y}) \mapsto \langle \vec{x},\vec{y}\rangle$ ahí definida, donde $\vec{x} = (x_1,\ldots,x_n)$ y $\vec{y} = (y_1,\ldots,y_n)$, es un producto interior en el espacio $\mathbb{R}^n$ indicado.

5 $\langle \vec{x},\vec{y}\rangle = 2x_1y_1 + x_2y_1 + x_1y_2 + 5x_2y_2$ en $\mathbb{R}^2$.

6 $\langle \vec{x},\vec{y}\rangle = 5y_1x_1 - 11y_1x_2 - 11y_2x_1 + 25y_2x_2$ en $\mathbb{R}^2$.

7 $\langle \vec{x},\vec{y}\rangle = x_1y_1 + 2x_2y_1 + 2x_1y_2 + 5x_2y_2$ en $\mathbb{R}^2$.

8 $\langle \vec{x},\vec{y}\rangle = 5x_1y_1 - x_2y_1 - x_1y_2 + x_2y_2$ en $\mathbb{R}^2$.

9 $\langle \vec{x},\vec{y}\rangle = 4x_1y_1 - 2x_2y_1 - 2x_1y_2 + 2x_2y_2$ en $\mathbb{R}^2$.

10 $\langle \vec{x},\vec{y}\rangle = 3x_1y_1 - 2x_2y_1 - 2x_1y_2 + 2x_2y_2$ en $\mathbb{R}^2$.

11 $\langle \vec{x},\vec{y}\rangle = x_1y_1 + 4x_2y_2 + x_3y_3$ en $\mathbb{R}^3$.

12 $\langle \vec{x},\vec{y}\rangle = x_1y_1 + 5x_2y_2 + x_3y_2 + x_2y_3 + x_3y_3$ en $\mathbb{R}^3$.

13 $\langle \vec{x},\vec{y}\rangle = 2x_1y_1 + x_2y_1 + x_1y_2 + x_2y_2 + x_3y_3$ en $\mathbb{R}^3$.

14 $\langle \vec{x},\vec{y}\rangle = x_1y_1 + x_3y_1 + x_2y_2 + x_1y_3 + 2x_3y_3$ en $\mathbb{R}^3$.

15 $\langle \vec{x},\vec{y}\rangle = 6x_1y_1 - x_1y_2 - x_2y_1 + 6x_2y_2 + x_2y_3 + x_3y_2 + 2x_3y_3$ en $\mathbb{R}^3$.

16 $\langle \vec{x},\vec{y}\rangle = 2x_1y_1 - x_3y_1 + x_2y_2 + x_3y_2 - x_1y_3 + x_2y_3 + 2x_3y_3$ en $\mathbb{R}^3$.

17 $\langle \vec{x},\vec{y}\rangle = x_1y_1 + x_2y_2 + 5x_3y_3 - x_4y_3 - x_3y_4 + 2x_4y_4$ en $\mathbb{R}^4$.

18 $\langle \vec{x},\vec{y}\rangle = x_1y_1 + 5x_2y_2 - 3x_3y_2 - 3x_2y_3 + 2x_3y_3 + x_4y_4$ en $\mathbb{R}^4$.

19 $\langle \vec{x},\vec{y}\rangle = 2x_1y_1 + 2x_2y_2 + x_3y_3 - x_4y_3 - x_3y_4 + 2x_4y_4 + x_5y_5$ en $\mathbb{R}^5$.

20 Sean $\langle \vec{x},\vec{y}\rangle$ el producto interior dado en el ejercicio 5, $\|\cdot\|$ la norma inducida por este producto (i.e., $\|\vec{x}\| = \langle \vec{x},\vec{y}\rangle^{1/2}$), $\vec{x} = (-1,2)$, $\vec{y} = (1,-1)$ y $\vec{w} = (-1,3)$; calcular:

(a) $\langle\vec{x},\vec{w}\rangle$.

(b) $\|\vec{w}\|$.

(c) $\mathrm{d}(\vec{x},\vec{y})$, la distancia entre $\vec{x}$ y $\vec{y}$ con la norma inducida.

21 Sean $\langle\vec{x},\vec{y}\rangle$ el producto interior dado en el ejercicio 6, $\|\cdot\|$ la norma inducida por este producto (i.e., $\|\vec{x}\| = \langle\vec{x},\vec{y}\rangle^{1/2}$), $\vec{x} = (-3,2)$, $\vec{y} = (-1,2)$ y $\vec{w} = (-1,-3)$; calcular:

(a) $\langle\vec{x},\vec{w}\rangle$.

(b) $\|\vec{w}\|$.

(c) $\mathrm{d}(\vec{x},\vec{y})$, la distancia entre $\vec{x}$ y $\vec{y}$ con la norma inducida.

22 Sean $\langle\vec{x},\vec{y}\rangle$ el producto interior dado en el ejercicio 12, $\|\cdot\|$ la norma inducida por este producto (i.e., $\|\vec{x}\| = \langle\vec{x},\vec{y}\rangle^{1/2}$), $\vec{x} = (1,-2,1)$, $\vec{y} = (2,-2,-1)$ y $\vec{w} = (1,1,-3)$; calcular:

(a) $\langle\vec{x},\vec{w}\rangle$.

(b) $\|\vec{w}\|$.

(c) $\mathrm{d}(\vec{x},\vec{w})$, la distancia entre $\vec{x}$ y $\vec{w}$ con la norma inducida.

23 Sean $\langle\vec{x},\vec{y}\rangle$ el producto interior dado en el ejercicio 19, $\|\cdot\|$ la norma inducida por este producto (i.e., $\|\vec{x}\| = \langle\vec{x},\vec{y}\rangle^{1/2}$), $\vec{x} = (1,-1,3,-1,1)$, $\vec{y} = (2,1,-1,0,1)$ y $\vec{w} = (1,0,-3,-1,2)$; calcular:

(a) $\langle\vec{x},\vec{w}\rangle$.

(b) $\|\vec{w}\|$.

(c) $\mathrm{d}(\vec{x},\vec{w})$, la distancia entre $\vec{x}$ y $\vec{w}$ con la norma inducida.

En los ejercicios 24 a 29 utilizar el ejercicio resuelto 10 de esta sección para encontrar la representación matricial del producto interior dado (cfr. ejercicio resuelto 11).

24 Del producto interior del ejercicio propuesto 5.

25 Del producto interior del ejercicio propuesto 6.

26 Del producto interior del ejercicio propuesto 8.

27 Del producto interior del ejercicio propuesto 12.

28 Del producto interior del ejercicio propuesto 14.

29 Del producto interior del ejercicio propuesto 16.

30 Sea A una matriz simétrica de orden n.

(a) Se define $\langle\vec{x},\vec{y}\rangle = \vec{x}^{\,t}A\vec{y}$, para cada $\vec{x},\vec{y}\in\mathbb{R}^n$. Probar que $\langle\vec{x},\vec{y}\rangle$ satisface las primeras tres condiciones de la definición 4.1 de producto interior.

(b) Dar un ejemplo de una matriz cuadrada simétrica de orden $n = 2$ que no cumpla con las últimas dos condiciones de la definición 4.1 de la página 236.

(*c*) La matriz A se define positiva si $\vec{x}^t A\vec{x} > 0$ para todo $\vec{x} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$. Dar un ejemplo de una matriz simétrica A de orden $n = 2$ que sea definida positiva.

(*d*) Probar que si A es una matriz simétrica definida positiva, entonces $\langle \vec{x}, \vec{y} \rangle = \vec{x}^t A\vec{y}$ es un producto interior en $\mathbb{R}^n$.

31 Sea A una matriz simétrica de orden $n = 2$, $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$.

(*a*) Si $a > 0$, probar que A es definida positiva (cfr. ejercicio precedente) si y sólo si $ad - b^2 > 0$.

(*b*) Si $a < 0$, mostrar que A no es definida positiva.

En los ejercicios 32 a 37, $\vec{x} = (x_1, \ldots, x_n)$ y $\vec{y} = (y_1, \ldots, y_n)$; determinar si $\langle \vec{x}, \vec{y} \rangle$, definido por la fórmula dada, es o no un producto interior. En caso afirmativo, demostrar rigurosamente que se cumplen las 5 condiciones de la definición 4.1, y en caso negativo mostrar, mediante contraejemplos, las propiedades que no se satisfacen.

32 $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i |y_i|$.

33 $\langle \vec{x}, \vec{y} \rangle = \left(\sum_{i=1}^n x_i^2 y_i^2\right)^{1/2}$.

34 $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n |x_i y_i|$.

35 $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i \sum_{i=1}^n y_i$.

36 $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n (x_i + y_i)^2 - \sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2$.

37 $\langle \vec{x}, \vec{y} \rangle = |\sum_{i=1}^n x_i y_i|$.

En los ejercicios 38 a 59 $\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \sum_{i=1}^n x_i y_i$ es el producto interior canónico en los espacios $\mathbb{R}^n$; $\vec{u} = (-1,-1,2)$, $\vec{v} = (-1,2,1)$, $\vec{w} = (-1,-1,1)$; $\vec{u}_1 = (2,-1,1,1)$, $\vec{v}_1 = (-1,0,1,0)$, $\vec{w}_1 = (1,-2,1,0)$ y $\vec{z}_1 = (1,-2,1,0)$

38 Calcular $\langle \vec{u}, \vec{v} \rangle$.

39 Hallar $\langle \vec{u}, v - \vec{w} \rangle$.

40 Encontrar $\langle 2\vec{w} - \vec{v}, 3\vec{u} - 2\vec{v} \rangle$.

41 Calcular $\|\vec{w}\|$.

42 Hallar $\mathrm{d}(\vec{u}, \vec{w})$.

43 Estimar el ángulo entre $\vec{u}$ y $\vec{v}$.

44 Hallar el vector proyección de $\vec{u}$ sobre el vector $\vec{w}$.

45 Encontrar el vector proyección del vector $\vec{u}$ sobre el subespacio $S = \mathrm{gn}(\vec{v}, \vec{w})$.

46 Hallar la matriz de proyección para el subespacio $S = \mathrm{gn}(\vec{u}, \vec{w})$.

47 Encontrar la proyección del vector $\vec{v}$ sobre el subespacio $S = \text{gn}(\vec{u}, \vec{w})$.

48 Hallar la proyección del vector $\vec{u}$ sobre el plano $x_1 - 2x_2 + x_3 = 0$.

49 Encontrar la proyección del vector $\vec{v}$ sobre el plano $2x_1 - x_2 + 3x_3 = 0$.

50 Calcular $\langle \vec{w}_1, \vec{v}_1 \rangle$.

51 Hallar $\langle 3\vec{z}_1, \vec{v}_1 - \vec{u}_1 \rangle$.

52 Encontrar $\langle 2\vec{w}_1 - 3\vec{v}_1, \vec{u}_1 - 2\vec{z}_1 \rangle$.

53 Calcular $\|\vec{z}_1\|$.

54 Hallar $\text{d}(\vec{z}_1, \vec{w}_1)$.

55 Estimar el ángulo entre $\vec{u}_1$ y $\vec{w}_1$.

56 Hallar el vector proyección de $\vec{v}_1$ sobre el vector $\vec{w}_1$.

57 Encontrar el vector proyección para el subespacio $S = \text{gn}(\vec{v}_1, \vec{w}_1)$.

58 Hallar la matriz de proyección para el subespacio $S = \text{gn}(\vec{u}_1, \vec{w}_1, \vec{z}_1)$.

59 Encontrar la proyección del vector $\vec{v}_1$ sobre el subespacio $S = \text{gn}(\vec{u}_1, \vec{w}_1, \vec{z}_1)$.

En los ejercicios 60 a 72 $\langle f, g \rangle = \int_a^b f(x)g(x)dx$ es el producto interior en el espacio $C[-1,1]$ (cfr. ejemplo 4.3, pág. 237) y $\|\cdot\|$ es la norma inducida correspondiente a este producto. Encontrar el valor indicado.

60 $\langle 1, x \rangle$.

61 $\langle x, 1+x \rangle$.

62 $\langle x, e^x \rangle$.

63 $\|x\|$.

64 $\|e^x\|$.

65 $\|1+x\|$.

66 $\text{d}(x, x^2)$.

67 $\text{d}(x, e^x)$.

68 El ángulo entre $f(x) = 1$ y $g(x) = x$.

69 El ángulo entre $f(x) = x$ y $g(x) = e^x$.

70 La proyección de $f(x) = e^x$ sobre la función $g(x) = x$.

71 La proyección de $f(x) = x$ sobre $g(x) = 1$.

72 La proyección de $f(x) = e^x$ sobre el subespacio $S = \text{gn}(1, x)$.

En los ejercicios 73 a 82 $\langle A, B \rangle = \text{tra}(B^t A)$ es el producto interior en el espacio de matrices (cfr. ejemplos 4.6 y 4.8, págs. 240 y 242) y $\|\cdot\|$ es la norma inducida correspondiente a este producto. Encontrar el valor indicado.

73 $\left\langle \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \right\rangle$.

74 $\left\| \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} \right\|$.

75 $\left\langle \begin{bmatrix} 2 & -1 & 1 \\ 1 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \right\rangle$.

76 $\left\| \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & -3 \end{bmatrix} \right\|$.

77 El ángulo entre $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$.

78 El ángulo entre $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 0 & 3 \end{bmatrix}$.

79 La proyección de $A = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$ sobre $B = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

80 La proyección de $A = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 1 & 0 \end{bmatrix}$ sobre $B = \begin{bmatrix} -1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

81 La proyección de $A = \begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix}$ sobre el subespacio

$$S = \text{gn}\left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \right).$$

82 La proyección de $A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & -1 & 2 \end{bmatrix}$ sobre el subespacio

$$S = \text{gn}\left(\begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 & 1 \\ 1 & -1 & 2 \end{bmatrix} \right).$$

En los ejercicios 83 a 94 $\langle (a_n),(b_n)\rangle = \sum_{n=1}^{\infty} a_n b_n$ es el producto interior en el espacio ℓ_2 de sucesiones de cuadrado sumable (cfr. ejemplo 4.9, pág. 243) y $\|\cdot\|$ es la norma inducida correspondiente a este producto. Probar que las sucesiones dadas pertenecen a ℓ_2 y encontrar los valores indicados.

83 $A = \left(\frac{1}{n+1}\right)$, $B = \left(\frac{1}{n+2}\right)$; $\langle A,B\rangle$.

84 $A = \left(\frac{1}{2n+1}\right)$, $B = \left(\frac{1}{2n-1}\right)$; $\langle A,B\rangle$.

85 $A = \left(\frac{1}{2^n}\right)$, $\left(\frac{1}{n!}\right)$; $\langle A,B\rangle$.

86 $A = \left((-\frac{\pi^2}{16})^n\right)$, $B = \left((\frac{1}{(2n)!})\right)$; $\langle A,B\rangle$

87 $A = \left(\frac{1}{5^n}\right)$, $B = \left(\left(-\frac{1}{\sqrt{3}}\right)^n\right)$; $\langle A,B\rangle$, $\|A\|$, $\|B\|$ y el ángulo entre A y B.

88 $A = \left((-\frac{1}{\sqrt{5}})^n\right)$; $\|A\|$.

89 $A = \left(\frac{1}{3n+1}\right)$, $B = \left(\frac{1}{3n-2}\right)$; $\langle A,B\rangle$; $\|A\|$, $\|B\|$ y el ángulo entre A y B.

90 $A = \left(\frac{1}{4n+3}\right)$, $B = \left(\frac{1}{4n-1}\right)$; $\langle A,B\rangle$; $\|A\|$, $\|B\|$ y el ángulo entre A y B.

91 $A = \left(2^n 5^{-n+2}\right)$, $B = \left(3^{n-2}4^{n-3}\right)$; $\langle A,B\rangle$, $\|A\|$, $\|B\|$ y el ángulo entre A y B.

92 $A = \left(\frac{1}{3^n}\right)$, $B = \left((-\frac{1}{4})^n\right)$; la proyección de A sobre B.

93 $A = \left(\frac{1}{2^n}\right)$, $B = \left(\frac{2}{3^n}\right)$, $C = \left((-\frac{1}{5})^n\right)$; la proyección de C sobre $\text{gn}(A,B)$.

94 $A = \left((-\frac{1}{\sqrt{2}})^n\right)$, $B = \left(\left(\frac{1}{\sqrt{3}}\right)^n\right)$, $C = \left(\frac{1}{3^n}\right)$; la proyección de C sobre $\text{gn}(A,B)$.

95 (**Complemento ortogonal de un subespacio**). Sea $\mathbf{E}$ un espacio con producto interior $\langle\cdot,\cdot\rangle$ y S un subespacio de $\mathbf{E}$. Probar que

$$S^{\perp} = \{\vec{u} \in \mathbf{E} \,|\, \vec{u} \perp \vec{v} \quad \forall \vec{v} \in S\}$$

es un subespacio de $\mathbf{E}$ (el llamado complemento ortogonal se S).

En los ejercicios 96 a 101 se hace referencia al complemento ortogonal definido en el ejercicio precedente.

96 Probar que $S \subset (S^{\perp})^{\perp}$.

97 Dar un ejemplo en el que $(S^{\perp})^{\perp} \neq S$.

98 Probar que $\mathscr{L}(S) \subset (S^{\perp})^{\perp}$.

99 Probar que si $\mathbf{E}$ tiene dimensión finita, entonces $(S^{\perp})^{\perp} = S$.

100 Sean S_1 y S_2 un par de subespacios en un espacio con producto interior. Probar que:

(a) $(S_1+S_2)^\perp = S_1^\perp \cap S_2^\perp$.

(b) $(S_1 \cap S_2)^\perp = S_1^\perp + S_2^\perp$.

101 Probar que si $S < \mathbf{E}$ y $\mathbf{E}$ tiene dimensión finita, entonces

$$\mathbf{E} = S \oplus S^\perp$$

(cfr. ejercicio resuelto 27 del capítulo 3, pág. 187).

En los ejercicios 102 a 109 los espacios a los que se hace referencia están dotados de los productos interiores usuales como se han definido en este texto (ejemplos 4.2, 4.3, 4.6 y 4.8). En cada caso encontrar $S^\perp$, el complemento ortogonal del subespacio indicado S (cfr. ejercicio 95).

102 $S = \operatorname{gn}((-1,2,1),(2,-2,1))$ en $\mathbb{R}^3$.

103 $S = \operatorname{gn}((1,0,1),(-2,1,1))$ en $\mathbb{R}^3$.

104 $S = \operatorname{gn}((-2,1,1,1),(-3,2,1,1),(2,2,-2,3))$ en $\mathbb{R}^4$.

105 $S = \operatorname{gn}((1,-3,2,1),(2,-3,4,-5),(-1,-4,1,1))$ en $\mathbb{R}^4$.

106 $S = \operatorname{gn}\left(\begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & -3 \\ -2 & 4 \end{bmatrix}\right)$ en $\mathfrak{M}_2$.

107 $S = \operatorname{gn}\left(\begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}\right)$ en $\mathfrak{M}_2$.

108 En $\mathbf{P}_1$ con el producto interior $\langle f,g\rangle = \int_0^1 f(x)g(x)dx$, $S = \operatorname{gn}(1)$.

109 En $\mathbf{P}_2$ con el producto interior $\langle f,g\rangle = \int_0^1 f(x)g(x)dx$, $S = \operatorname{gn}(1,x)$.

110 Sean $C[-1,1]$ el espacio de funciones continuas dotado del producto interior usual

$$\langle f,g\rangle = \int_{-1}^1 f(x)g(x)dx,$$

S_P y S_I los subespacios de funciones pares e impares, respectivamente, probar que $S_P^\perp = S_I$ (cfr. ejercicio 95).

111 ¿Es cierta la conclusión del ejercicio anterior si se trabaja en el espacio $C[0,1]$? ¿Cómo deben ser a,b para que la conclusión del ejercicio precedente sea válida en $C[a,b]$ con el producto interior

$$\langle f,g\rangle = \int_a^b f(x)g(x)dx?$$

112 Sean $\vec{u}$ y $\vec{v}$ un par de vectores en un espacio con producto interior $\langle\cdot,\cdot\rangle$.

(a) Si $\vec{u} \perp \vec{v}$, calcular $\langle 3\vec{u} - 2\vec{v}, 2\vec{u} + \vec{v} \rangle$.

(b) Si $\|\vec{u}\| = 2$ y $\|\vec{v}\| = 7$, calcular $\langle \vec{u} - 4\vec{v}, \vec{u} + 4\vec{v} \rangle$.

113 Sea **E** un espacio con producto interior $\langle \cdot, \cdot \rangle$. Demostrar el recíproco del teorema de Pitágoras; i.e., si $\vec{u}, \vec{v} \in \mathbf{E}$ y

$$\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2,$$

entonces $\vec{u} \perp \vec{v}$.

114 Probar que en cualquier espacio con producto interior:

(a) $\|\vec{u} - \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$.

(b) $|\|\vec{u}\| - \|\vec{v}\|| \leq \|\vec{u} - \vec{v}\|$.

Para todo par de vectores $\vec{u}, \vec{v} \in \mathbf{E}$.

115 Probar que en un espacio con producto interior los vectores $\|\vec{u}\|\,\vec{v} + \|\vec{v}\|\,\vec{u}$ y $\|\vec{u}\|\,\vec{v} - \|\vec{v}\|\,\vec{u}$ son ortogonales.

En los ejercicios 116 a 119 **E** es un espacio con producto interior y $\vec{u}$ y $\vec{v}$ son un par de vectores en él. Probar la afirmación dada.

116 $\vec{u} \perp \vec{v} \Leftrightarrow \|\vec{u} + \vec{v}\| = \|\vec{u} - \vec{v}\|$.

117 $\vec{u} + \vec{v} \perp \vec{u} - \vec{v} \Leftrightarrow \|\vec{u}\| = \|\vec{v}\|$.

118 $\vec{u} \perp \vec{v} \Rightarrow \|\vec{u} + \alpha\vec{v}\| \geq \|\vec{u}\|$ para todo $\alpha \in \mathbb{R}$.

119 Si $\vec{u}, \vec{v} \neq \vec{0}_{\mathbf{E}}$, entonces

$$\|\vec{u} - \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\,\|\vec{u}\|\,\|\vec{v}\|\cos\theta.$$

120 Sea **E** un espacio vectorial y $\langle \cdot, \cdot \rangle_1$, $\langle \cdot, \cdot \rangle_2$ un par de productos interiores en él. Se define, $\langle \vec{u}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle_1 + \langle \vec{u}, \vec{v} \rangle_2$. ¿Es $\langle \vec{u}, \vec{v} \rangle$ un producto interior en **E**?

En los ejercicios 121 a 145 los espacios considerados están dotados de los productos interiores usuales definidos en los ejemplos 4.2, 4.3, 4.6, 4.8 y 4.9 de este texto; encontrar una base ortonormal para el subespacio generado por el conjunto L.I. dado o para el subespacio indicado utilizando el proceso de ortogonalización de Gram-Schmidt (cfr. teorema 4.12, pág. 275).

121 $\{(-1,1),(1,2)\}$ en $\mathbb{R}^2$.

122 $\{(2,2),(1,-1)\}$ en $\mathbb{R}^2$.

123 $\{(0,2),(3,1)\}$ en $\mathbb{R}^2$.

124 $\{(1,1),(0,1)\}$ en $\mathbb{R}^2$.

125 $\{(3,1,1),(-1,2,2)\}$ en $\mathbb{R}^3$.

126 $\{(0,1,1),(1,2,-1)\}$ en $\mathbb{R}^3$.

127 $\{(-1,1,0),(2,-1,2),(1,1,1)\}$ en $\mathbb{R}^3$.

128 $\{(2,1,-1),(1,-1,1),(-1,1,1)\}$ en $\mathbb{R}^3$.

129 $\{(-1,1,0),(2,-1,1),(1,1,1)\}$ en $\mathbb{R}^3$.

130 $\{(3,1,-3),(1,-1,2),(-1,1,-1)\}$ en $\mathbb{R}^3$.

131 $\{(1,0,1,0),(1,1,1,0),(1,-1,0,1)\}$ en $\mathbb{R}^4$.

132 $\{(1,-1,1,0,0),(-1,0,0,0,1),(0,0,1,0,1),(1,0,0,1,1)\}$ en $\mathbb{R}^5$.

133 $\{\operatorname{sen} x, \cos x\}$ en $C[0,\pi]$.

134 $\{1,x,x^2\}$ en $C[-1,1]$.

135 $\{1,x,x^2\}$ en $C[0,1]$.

136 $\{1,e^x\}$ en $C[-1,1]$.

137 $\left\{\begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\right\}$ en $\mathfrak{M}_2$.

138 $\left\{\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ en $\mathfrak{M}_2$.

139 $\left\{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}\right\}$ en $\mathfrak{M}_{2\times 3}$.

140 El plano $x-2y+z=0$ en $\mathbb{R}^3$.

141 El espacio solución del sistema

$$\begin{aligned} x-y+z-w&=0\\ 2x+y-4z+2w&=0\\ 3x-y+4z+w&=0 \end{aligned}$$

en $\mathbb{R}^4$.

142 El espacio nulo de la matriz

$$\begin{bmatrix} 1 & -1 & 2 & 1 \\ -1 & 2 & -2 & 3 \\ 2 & -1 & 0 & 0 \\ 1 & 1 & -2 & 3 \end{bmatrix}$$

en $\mathbb{R}^4$.

143 El espacio nulo de la matriz

$$\begin{bmatrix} -1 & 1 & 3 \\ 2 & 1 & -5 \\ -1 & 4 & 4 \end{bmatrix}$$

en $\mathbb{R}^3$.

144 $\left\{\left(\frac{1}{2^n}\right), \left(\frac{1}{3^n}\right)\right\}$ en ℓ_2.

145 $\left\{\left(\frac{1}{2^n}\right), \left(\frac{1}{3^n}\right), \left(\frac{1}{4^n}\right)\right\}$ en ℓ_2.

En los ejercicios 146 a 159 $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$ es el producto interior usual en el espacio $C[-1,1]$, $\|\cdot\|$ la norma inducida correspondiente, $f(x) = 1$, $g(x) = x$, $h(x) = 1+x$ y $r(x) = x^2$.

146 Calcular $\langle f, g\rangle$.

147 Hallar $\langle 1, h\rangle$.

148 Calcular $\langle g, r\rangle$.

149 Encontrar $\|g\|$.

150 Calcular $\|h\|$.

151 Hallar $\|r\|$

152 Hallar el ángulo entre f y g.

153 Calcular el ángulo entre f y h.

154 Hallar el ángulo entre g y h.

155 Encontrar la proyección de r sobre f.

156 Encontrar la proyección de g sobre r.

157 Hallar la proyección de r sobre el subespacio $S = \mathrm{gn}(1, x)$.

158 Encontrar todos los vectores en $S = \mathrm{gn}(1, x)$ que son ortogonales a f.

159 Hallar una base ortonormal para el espacio $S = \mathrm{gn}(1, x, x^2)$.

En los ejercicios 160 a 167 $\langle f, g\rangle$ es el producto interior definido en el ejercicio resuelto 4 de esta sección.

160 Si $f_n(x) = x^n$, $n = 0, 1, 2, \ldots$, probar que $\langle f_m, f_n\rangle = (m+n)!$

161 Si $f(x) = (1+x)^2$ y $g(x) = x^2 + 1$, calcular $\langle f, g\rangle$.

162 Encontrar todos los polinomios en $\mathrm{gn}(1,x)$ que son ortogonales a $f(x)=1+x$.

163 Hallar el ángulo entre $f(x)=x$ y $g(x)=x^2$.

164 Encontrar la proyección de $f(x)=x$ sobre $g(x)=x^2$.

165 Encontrar la proyección de $g(x)=x^2$ sobre el subespacio $S=\mathrm{gn}(1,x)$.

166 Hallar una base ortonormal para el subespacio $S=\mathrm{gn}(1,x)$.

167 Encontrar una base ortonormal para el subespacio $S=\mathrm{gn}(1,x,x^2)$.

168 Sea $C[1,e]$ el espacio de funciones continuas en $[1,e]$, se define

$$\langle f,g\rangle=\int_1^e f(x)g(x)\ln(x)dx$$

para cada $f,g\in C[1,e]$. Probar que $\langle f,g\rangle$ es un producto interior en $C[1,e]$.

En los ejercicios 169 a 179 $\langle f,g\rangle=\int_1^e f(x)g(x)\ln(x)dx$ es el producto interior definido en el ejercicio precedente en el espacio $C[1,e]$, $\|\cdot\|$ la norma inducida por este producto, $f(x)=1$, $g(x)=x$, $h(x)=x^2$ y $r(x)=\sqrt{x}$.

169 Calcular $\langle f,g\rangle$.

170 Hallar $\langle 1,h\rangle$.

171 Calcular $\langle g,r\rangle$.

172 Encontrar $\|g\|$.

173 Calcular $\|h\|$.

174 Hallar $\|r\|$

175 Encontrar la proyección de r sobre f.

176 Encontrar la proyección de g sobre r.

177 Hallar la proyección de r sobre el subespacio $S=\mathrm{gn}(1,x)$.

178 Encontrar todos los vectores en $S=\mathrm{gn}(1,x)$ que son ortogonales a f.

179 Hallar una base ortonormal para el espacio $S=\mathrm{gn}(1,x,x^2)$.

180 En el espacio de polinomios de grado a lo más n, $\mathbf{P}_n$, se define, para cada par de polinomios f y g,

$$\langle f,g\rangle=\sum_{k=0}^{n} f\left(\frac{k}{n}\right)g\left(\frac{k}{n}\right).$$

Probar que $\langle f,g\rangle$ es un producto interior en $\mathbf{P}_n$.

En los ejercicios 181 a 186 $\langle f,g\rangle = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) g\left(\frac{k}{n}\right)$ es el producto interior en el espacio $\mathbf{P}_n$ dado en el ejercicio anterior y $\|\cdot\|$ la norma inducida por este producto.

181 Calcular $\langle f,g\rangle$ si $f(x) = x$ y $g(x) = a_0 + a_1 x$.

182 Calcular $\|g(x)\|$ si $g(x) = 1$.

183 Calcular $\|f\|$ si $f(x) = x$.

184 Encontrar el ángulo entre $f(x) = 1$ y $g(x) = x$.

185 Encontrar todos los polinomios de grado a lo más uno que son ortogonales a $f(x) = x$.

186 Encontrar una base ortonormal para el subespacio $S = \text{gn}(1,x)$.

En los ejercicios 187 a 189 **E** es el conjunto de todas las funciones continuas $f : [0,\infty) \to \mathbb{R}$ tales que

$$\int_0^\infty [f(x)]^2 e^{-x} dx$$

converge.

187 Probar que

$$\int_0^\infty f(x)g(x)e^{-x} dx$$

converge absolutamente (i.e., $\int_0^\infty |f(x)g(x)| e^{-x} dx$ converge) para todo par de funciones $f,g \in \mathbf{E}$.

188 Mostrar que **E**, con la suma de funciones y producto de un escalar por una función usuales, es un espacio vectorial.

189 Probar que $\langle f,g\rangle = \int_0^\infty f(x)g(x)e^{-x} dx$ define un producto interior en el espacio vectorial **E**.

En los ejercicios 190 a 210 $\langle f,g\rangle = \int_0^\infty f(x)g(x)e^{-x} dx$ es el producto interior en el espacio **E** definido en los ejercicios 187 a 189 y $\|\cdot\|$ es la norma inducida por este producto.

190 Probar que toda función de la forma $f_n(x) = x^n$, $n = 0,1,2,\ldots$, pertenece a **E**.

191 ¿La función $f(x) = e^x$ pertenece a **E**?

192 Probar que $f(x) = e^{-x}$ pertenece a **E**.

193 Encontrar $\langle 1,x\rangle$.

194 Calcular $\|1\|$.

195 Hallar $\|x\|$.

196 Hallar la proyección de $f(x) = x$ sobre $g(x) = 1$.

197 Hallar una base ortonormal para el subespacio $S = \text{gn}(1, x)$.

198 Encontrar una base ortonormal para el subespacio $S = \text{gn}(1, x, x^2)$.

199 Si $f(x) = x^n$, $n = 0, 1, \ldots,$ y $g(x) = e^{-x}$, calcular $\langle f, g \rangle$.

200 Encontrar el ángulo entre $f(x) = 1$ y $g(x) = e^{-x}$.

201 Calcular $\|e^{-x}\|$.

202 Encontrar la proyección de $f(x) = e^{-x}$ sobre $g(x) = 1$.

203 Encontrar la proyección de $f(x) = e^{-x}$ sobre $S = \text{gn}(1, x)$.

204 Probar que $f(x) = \cos x$ y $g(x) = \text{sen}(x)$ pertenecen a $\mathbf{E}$.

205 Calcular $\langle \cos x, e^{-x} \rangle$ y $\langle \text{sen}\, x, e^{-x} \rangle$.

206 Hallar $\|\cos x\|$.

207 Calcular $\|\text{sen}\, x\|$.

208 Encontrar la proyección de $f(x) = e^{-x}$ sobre $g(x) = \cos x$.

209 Hallar una base ortonormal para el subespacio $S = \text{gn}(\cos x, \text{sen}\, x)$.

210 Encontrar la proyección de $f(x) = e^{-x}$ sobre $S = \text{gn}(\cos x, \text{sen}\, x)$.

En los ejercicios 211 a 214: *(i)* probar que las columnas de la matriz A son L.I. *(ii)* Encontrar una factorización QR para la matriz A; i.e., una matriz Q cuyas columnas son ortonormales y una matriz R que es triangular superior e invertible tales que $A = QR$ (cfr. teorema 4.13, pág. 279).

211 $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & 0 \\ -1 & 3 & 1 \\ 1 & 2 & -1 \end{bmatrix}$.

212 $A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

213 $A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$.

214 $A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

En los ejercicios 215 a 221, determinar si la matriz dada es o no ortogonal (cfr. teorema 4.14, pág. 282).

215 $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

216 $\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$.

217 $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

218 $\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$.

219 $\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$.

220 $\frac{1}{5}\begin{bmatrix} -3 & 0 & 4 \\ 4 & 0 & 3 \\ 0 & -5 & 0 \end{bmatrix}$.

221 $\begin{bmatrix} \frac{1}{3} & 0 & \frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 0 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & 0 & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$.

En los ejercicios 222 a 228, $A \in \mathfrak{M}_n$ es un matriz ortogonal y $\vec{x}, \vec{y} \in \mathbb{R}^n$.

222 Probar que $A\vec{x} \cdot A\vec{y} = \vec{x} \cdot \vec{y}$.

223 Mostrar que $||A\vec{x}|| = ||\vec{x}||$.

224 Demostrar que el ángulo entre $A\vec{x}$ y $A\vec{y}$ es el mismo que entre $\vec{x}$ y $\vec{y}$.

225 Probar que $||A\vec{x}|| = \left\|A^{-1}\vec{x}\right\|$.

226 Demostrar que A^2 también es una matriz ortogonal.

227 Probar que $\det(A) = \pm 1$.

228 Si $C \in \mathfrak{M}_n$ es tal que $D = A^{-1}CA$ es una matriz diagonal, mostrar que C es simétrica.

229 Dar un ejemplo de una matriz A tal que $\det(A) = 1$, pero A no es ortogonal.

230 Probar que si A es una matriz cuadrada de orden n y $A\vec{x}\cdot A\vec{y} = \vec{x}\cdot\vec{y}$ para todo par de vectores $\vec{x},\vec{y}\in\mathbb{R}^n$, entonces A es una matriz ortogonal.

231 Si A es una matriz cuadrada de orden n tal que $\|A\vec{x}\| = \|\vec{x}\|$ para todo $\vec{x}\in\mathbb{R}^n$, demostrar que A es una matriz ortogonal.

232 Probar que si A es una matriz cuadrada de orden n que es ortogonal y $\{\vec{u}_1,\ldots,\vec{u}_n\}$ es una base ortonormal de $\mathbb{R}^n$, entonces $\{A\vec{u}_1,\ldots,A\vec{u}_n\}$ es también una base ortonormal de este espacio.

En los ejercicios 233 a 241, $\mathbf{E}$ y $\langle f,g\rangle = \int_0^\infty f(x)g(x)e^{-x^2}dx$ son el espacio y el producto interior de los ejercicios resueltos 13 y 14 y $\|\cdot\|$ es la norma inducida por este producto escalar.

233 Calcular $\|\cos x\|$.

234 Probar que si $f(x) = \operatorname{sen} x$, entonces $f\in\mathbf{E}$ y calcular $\|f\|$.

235 Probar que si $f(x) = \cos x$, entonces $f\in\mathbf{E}$ y calcular $\|f\|$.

236 Calcular $\langle x, \operatorname{sen} x\rangle$.

237 Si $f(x) = x^n$, mostrar que $f\in\mathbf{E}$ para todo $n = 0,1,\ldots$

238 Mostrar que todo polinomio pertenece a $\mathbf{E}$.

239 Calcular $\left\langle x^2, \cos x\right\rangle$.

240 Hallar la aproximación óptima de $f(x) = \cos x$ en el subespacio $S = \operatorname{gn}(1,x^2)$.

241 Hallar la aproximación óptima de $f(x) = \operatorname{sen} x$ en el subespacio $S = \operatorname{gn}(x,x^3)$.

En los ejercicios 242 a 255, encontrar la aproximación óptima ($\vec{p}^*$) del elemento $\vec{u}$ en el subespacio S del espacio $\mathbf{E}$, con el producto interior $\langle\cdot,\cdot\rangle$ indicado (cfr. apartado 4.1.5, pág. 283).

242 $\mathbf{E} = \mathbb{R}^3$, $\langle\vec{x},\vec{y}\rangle = \vec{x}\cdot\vec{y}$, $S = \operatorname{gn}((-1,1,2),(-2,1,1))$, $\vec{u} = (1,1,1)$.

243 $\mathbf{E} = \mathbb{R}^3$, $\langle\vec{x},\vec{y}\rangle = \vec{x}\cdot\vec{y}$, $S = \operatorname{gn}((0,1,1),(-1,1,-1))$, $\vec{u} = (1,0,1)$.

244 $\mathbf{E} = \mathbb{R}^4$, $\langle\vec{x},\vec{y}\rangle = \vec{x}\cdot\vec{y}$, $S = \operatorname{gn}((1,1,2,1),(-1,1,1,0),(2,-1,1,3))$, $\vec{u} = (1,-2,4,1)$.

245 $\mathbf{E} = \mathbb{R}^4$, $\langle\vec{x},\vec{y}\rangle = \vec{x}\cdot\vec{y}$, $S = \operatorname{gn}((1,0,1,1),(1,-1,-1,0),(1,1,1,1))$, $\vec{u} = (0,-1,2,-1)$.

246 $\mathbf{E} = C[1,3]$, $\langle f,g\rangle = \int_1^3 f(x)g(x)dx$, $S = \operatorname{gn}(1)$, $\vec{u} = f$; donde $f(x) = 1/x$.

247 $\mathbf{E} = C[0,1]$, $\langle f,g\rangle = \int_0^1 f(x)g(x)dx$, $S = \operatorname{gn}(1,x)$, $\vec{u} = f$; donde $f(x) = \frac{1}{x-2}$.

248 $\mathbf{E} = C[0,1]$, $\langle f,g\rangle = \int_0^1 f(x)g(x)dx$, $S = \text{gn}(1,x)$, $\vec{u} = f$; donde $f(x) = e^{-x}$.

249 $\mathbf{E} = C[0,1]$, $\langle f,g\rangle = \int_0^1 f(x)g(x)dx$, $S = \text{gn}(1,x,x^2)$, $\vec{u} = f$; donde $f(x) = e^{-x}$.

250 $\mathbf{E} = C[0,2]$, $\langle f,g\rangle = \int_0^2 f(x)g(x)dx$, $S = \text{gn}(1)$, $\vec{u} = f$; donde $f(x) = e^{x}$.

251 $\mathbf{E} = C[-1,1]$, $\langle f,g\rangle = \int_{-1}^1 f(x)g(x)dx$, $S = \text{gn}(1,x)$, $\vec{u} = f$; donde $f(x) = e^{x}$.

252 $\mathbf{E} = C[0,2\pi]$, $\langle f,g\rangle = \int_0^{2\pi} f(x)g(x)dx$, $S = \text{gn}(1,\cos x,\operatorname{sen} x)$, $\vec{u} = f$; donde $f(x) = x$.

253 $\mathbf{E}$ es el espacio del ejercicio 187, $\langle f,g\rangle = \int_0^{\infty} f(x)g(x)e^{-x}dx$, $S = \text{gn}(1,x)$, $\vec{u} = f$; donde $f(x) = e^{-x}$.

254 $\mathbf{E}$ es el espacio del ejercicio resuelto 4, $\langle f,g\rangle = \int_0^{\infty} f(x)g(x)e^{-x}dx$, $S = \text{gn}(1,x)$, $\vec{u} = f$; donde $f(x) = x^2$.

255 $\mathbf{E} = \ell_2$, $\langle f,g\rangle = \sum_{n=0}^{\infty} a_n b_n$, $S = \text{gn}((1/2^n),((1/3^n))$, $\vec{u} = ((1/\sqrt{2})^n)$.

En los ejercicios 256 a 263, ajustar por mínimos cuadrados el conjunto de datos (x_i, y_i) a un polinomio del grado n indicado.

256 ; $n = 1$.

x_i	0	1	2	3	4
y_i	−1.98	1.01	3.98	6.95	10.01

257 ; $n = 1$.

x_i	−1	−2	0	3	4
y_i	3.97	5.02	3.01	0	−1.02

258 ; $n = 1$.

x_i	−1	2	3	5	7
y_i	−0.97	5.01	6.98	11.03	14.99

259 ; $n = 1$.

x_i	−3	2	1	2	4
y_i	11.04	0.96	3.01	1.01	−3.02

260 ; $n = 2$.

x_i	0	1	2	3
y_i	1.02	3.99	8.97	16.01

261 ; $n = 2$.

x_i	−1	0	1	2
y_i	−1.02	1.97	3.02	1.98

262 ; $n = 2$.

x_i	0	1	2	3
y_i	0.02	−.01	2.03	6.01

263

x_i	1	2	3	4
y_i	1.02	6.95	16.97	31

; $n = 2$.

264 En el laboratorio de electricidad los estudiantes de la materia de circuitos eléctricos intentan encontrar una relación entre la variación de una resistencia con su temperatura. Han realizado las mediciones que se resumen en la siguiente tabla:

T °C	R (omhs)
20.22	765
32.65	826
52	872
73.3	941
95.6	1 032.3

(a) Graficar los datos de esta tabla para conjeturar que la relación entre la resistencia y la temperatura es lineal; i.e., $R = mT + b$ y ajustar estos datos por mínimos cuadrados para determinar m y b.

(b) Estimar la resistencia que corresponde a una temperatura de 90 °C.

(c) ¿Cuáles son las dimensiones de m y b?

265 La siguiente tabla contiene información respecto al número de pobladores, p, en diversos meses t, de cierta clase de roedores que se reproducen rápidamente en condiciones de laboratorio:

t_i	0	1	2	3	4	5	6	8	10
p_i	2	3	5	7	10	15	20	40	109

(a) Graficar estos datos para conjeturar que la relación entre p y t es exponencial; i.e.,

$$p = re^{st}.$$

(b) Poner $z_i = \ln p_i = \ln r + st_i$ y añadir esta nueva fila a la tabla precedente.

(c) Hacer un ajuste lineal por mínimos cuadrados a los datos (t_i, z_i) para calcular r y s.

(d) Estimar la población en el mes 12.

Espacios vectoriales normados (respuestas en páginas 1082-1084)

En los ejercicios 266 a 275 se consideran un espacio $\mathbb{R}^n$ o un espacio $\mathfrak{M}_{m\times n}$ provistos de las normas $\|\cdot\|_\infty$ o $\|\cdot\|_1$ tratados en los ejemplos 4.40, 4.41, 4.42, 4.43, 4.44 (cfr. págs. 303 a 306) y $d(\cdot,\cdot)$ la correspondiente distancia definida por cada norma, calcular:

266 $\|(-1,2,1,-2)\|_\infty$ en $(\mathbb{R}^4, \|\cdot\|_\infty)$.

267 $\|\lambda\vec{u}\|_\infty$, si $\vec{u} = (-1,2,1,4,2)$ y $\lambda = -\sqrt{2}$, en $(\mathbb{R}^5, \|\cdot\|_\infty)$.

268 $d(\vec{u}, v)$, si $\vec{u} = (2,-3,2,-5,1)$ y $\vec{v} = (-1,5,2,-3,0)$, en $(\mathbb{R}^5, \|\cdot\|_1)$.

269 $\|(-1,2,1,4,2)\|_1$ en $(\mathbb{R}^5, \|\cdot\|_1)$.

270 $d((-2,3,-3,4,-2),(-1,-1,2,-3-5))$ en $(\mathbb{R}^5, \|\cdot\|_1)$.

271 $\left\|\begin{bmatrix} -1 & 2 \\ -7 & -4 \end{bmatrix}\right\|_\infty$ en $(\mathfrak{M}_2, \|\cdot\|_\infty)$.

272 $\left\|\begin{bmatrix} -1 & 2 & 3 & 4 \\ -1 & 3 & 2 & 5 \\ -5 & -7 & 6 & 2 \end{bmatrix}\right\|_\infty$ en $(\mathfrak{M}_{3\times4}, \|\cdot\|_\infty)$.

273 $\left\|\begin{bmatrix} -3 & -1 & 2 & 3 \\ 2 & 3 & 6 & -4 \\ -4 & 5 & -8 & 3 \end{bmatrix}\right\|_1$ en $(\mathfrak{M}_{3\times4}, \|\cdot\|_1)$.

274 $d(A,B)$ en $(\mathfrak{M}_{2\times3}, \|\cdot\|_\infty)$ si $A = \begin{bmatrix} -1 & 2 & 3 \\ 2 & -3 & 5 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & -3 & 4 \\ 6 & 3 & -5 \end{bmatrix}$.

275 $d(A,B)$ en $(\mathfrak{M}_{3\times4}, \|\cdot\|_1)$ si

$$A = \begin{bmatrix} 1 & -2 & 0 & -1 \\ 0 & -1 & 2 & -2 \\ -1 & 2 & -2 & 3 \end{bmatrix} \text{ y } B = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & -2 & -3 & 1 \\ 2 & -3 & 1 & 1 \end{bmatrix}.$$

En los ejercicios 276 a 288, $\|\cdot\|_\infty$ es la norma uniforme en el espacio indicado $C[a,b]$; esto es, $\|f\|_\infty = \max\limits_{a\le x\le b} |f(x)|$ para cada $f \in C[a,b]$ (cfr. ejemplo 4.47, pág. 308). Calcular $\|f\|_\infty$ para cada función f en el espacio dado $C[a,b]$.

276 $f(x) = x^2 - 2x - 8$ en $C[-1,4]$.

277 $f(x) = 4 - 4x - x^2$ en $C[-4,4]$.

278 $f(x) = 3x - x^3$ en $C[-2,2]$.

279 $f(x) = x^4 - 2x^2$ en $C[-\sqrt{2}, \sqrt{2}]$.

280 $f(x) = x^{2/3} + 1$ en $C[0,8]$.

281 $f(x) = x - x^2$en $C[-4,2]$.

282 $f(x) = \sqrt{x^2+4}$ en $C[-2,1]$.

283 $f(x) = (x-1)^{1/3} + \frac{1}{2}(x+1)^{2/3}$en $C[-3,7]$.

284 $f(x) = 2 - e^x + 2x$ en $C[0,1]$.

285 $f(x) = 1 + e^{-\cos(x-1)}$ en $C[0,2]$.

286 $f(x) = \sum_{i=1}^{n} (x-i)^2$ en $C[1,n]$.

287 $f(x) = \sum_{i=1}^{n} |x - i|$ en $C[0,1]$.

288 $f(x) = \dfrac{x+1}{x^2+1}$ en $C[-2,4]$.

En los ejercicios 289 a 298, $\|f\|_1 = \int_a^b |f(x)|\,dx$ es la norma del ejemplo 4.46 (pág. 306) en el espacio $C[a,b]$ dado y $\mathrm{d}(\cdot,\cdot)$ es la distancia definida por esta norma. Calcular:

289 $\|f\|_1$ si $f(x) = x^2 - 2x + 2$ en $C[0,5]$.

290 $\|f\|_1$ si $f(x) = \ln x$ en $C[1/2,2]$.

291 $\|f\|_1$ si $f(x) = \cos 2x$ si $C[0,2\pi]$.

292 $\mathrm{d}(f,g)$ si $f(x) = \sqrt{x}$ y $g(x) = x^2$ en $C[0,1]$.

293 $\mathrm{d}(f,g)$ si $f(x) = 5x^2 - 6x$ y $g(x) = x^3$ en $C[-1,4]$.

294 $\mathrm{d}(f,g)$ si $f(x) = 2x - x^2$ y $g(x) = x^3$ en $C[-2,2]$.

295 $\mathrm{d}(f,g)$ si $f(x) = e^x$ y $g(x) = 1$ en $C[-1,1]$.

296 $\mathrm{d}(f,g)$ si $f(x) = \cos x$ y $g(x) = \mathrm{sen}\, x$ en $C[0,\pi/2]$.

297 $\|f\|_1$ si $f(x) = x^3 - x$ en $C[-2,1]$.

298 $\|f\|_1$ si $f(x) = \tan x$ en $C[-\frac{\pi}{4},\frac{\pi}{4}]$.

299 Sea $\mathscr{C}_0$ el conjunto de sucesiones reales (a_n) tales que $\mathrm{lím}_{n\to\infty} a_n = 0$.

(*a*) Probar que $\mathscr{C}_0 < \mathbb{R}^\infty$ y, por tanto, un espacio vectorial con las operaciones usuales de suma de sucesiones y producto de un escalar con una sucesión.

(*b*) Mostrar que para toda sucesión $(a_n) \in \mathscr{C}_0$ existe n_0 tal que $|a_n| \le |a_{n_0}|$ $\forall n \in \mathbb{N}$.

300 Sea $\mathscr{C}_0$ el espacio vectorial del ejercicio precedente. Si $A = (a_n)$ es cualquier sucesión en él, se define $\|A\| = \mathrm{máx}_{n\in\mathbb{N}} |a_n|$. Probar que $\|A\|$ es una norma en $\mathscr{C}_0$.

En los ejercicios 301 a 310, $\mathscr{C}_0$ es el espacio vectorial del ejercicio anterior provisto de la norma ahí definida:

$$\mathscr{C}_0 = \{(a_n) \in \mathbb{R}^\infty \mid \lim_{n\to\infty} a_n = 0\},$$

$$\|(a_n)\| = \max_{n\in\mathbb{N}} |a_n|.$$

Mostrar que la sucesión (a_n) pertenece a $\mathscr{C}_0$ y calcular $\|(a_n)\|$ si:

301 $a_n = \dfrac{1}{n}$.

302 $a_n = \dfrac{1}{n+2}$.

303 $a_n = e^{1/n} - 1$.

304 $a_n = \dfrac{n-1}{e^{(n-1)^2}}$.

305 $a_n = \dfrac{1}{n^2}$.

306 $a_n = \dfrac{\operatorname{sen}\left((2n-1)\frac{\pi}{2}\right)}{n}$.

307 $a_n = \dfrac{\ln n}{n}$.

308 $a_n = \dfrac{n-1}{n^2+1}$.

309 $a_n = \dfrac{2^{n-1}}{(n-1)!}$.

310 $a_n = \dfrac{1-\cos n}{n}$.

311 Sean $\|\cdot\|'$ y $\|\cdot\|''$ un par de normas sobre un mismo espacio vectorial $\mathbf{E}$, mostrar que $\|\vec{u}\| = \|\vec{u}\|' + \|\vec{u}\|''$ es también una norma en $\mathbf{E}$.

312 Sea n un número entero positivo y $C^n[a,b]$ el conjunto de funciones $f : [a,b] \to \mathbb{R}$ tales que f tiene derivada hasta el orden n con continuidad en el intervalo $[a,b]$ (la derivabilidad y continuidad en los extremos se considera por la derecha y por la izquierda, respectivamente). Se representa, para cada $k = 1,\ldots,n$, la k-ésima derivada de f por el símbolo $f^{(k)}$ y $f^{(0)}$ denota la función f.

(a) Probar que $C^n[a,b]$ es un subespacio de $C[a,b]$.

(b) Mostrar que

$$\|f\| = \text{máx}\left\{\|f\|_\infty, \|f'\|_\infty, \|f\|_\infty^{(2)}, \ldots, \left\|f^{(n)}\right\|_\infty\right\},$$

$f \in C^n[a,b]$, es una norma en $C^n[a,b]$, donde $\left\|f^{(k)}\right\|_\infty$ es la norma uniforme en el espacio $C^k[a,b]$, para cada $k = 0,1,\ldots,n$.

En los ejercicios 313 a 316, $C^n[a,b]$ y $\|f\|$ son el espacio y la norma definidos en el ejercicio precedente.

313 Determinar si $f(x) = xe^{-x}$ pertenece a $C^1[0,3]$; en caso afirmativo calcular $\|f\|$.

314 Determinar si $f(x) = \begin{cases} x^2 \ln x, \text{ si } x \neq 0 \\ 0, \text{ si } x = 0 \end{cases}$ pertenece a $C^1[0,3]$; en caso afirmativo calcular $\|f\|$.

315 Determinar si $f(x) = \begin{cases} x^3 \ln x, \text{ si } x \neq 0 \\ 0, \text{ si } x = 0 \end{cases}$ pertenece a $C^1[0,3]$; en caso afirmativo calcular $\|f\|$.

316 Determinar si $f(x) = \begin{cases} e^{-1/x^2}, \text{ si } x \neq 0 \\ 0, \text{ si } x = 0 \end{cases}$ pertenece a $C^1[-1,1]$; en caso afirmativo calcular $\|f\|$.

317 Sea $C^n[a,b]$ el espacio de funciones $f : [a,b] \to \mathbb{R}$ tales que la función f tiene derivada hasta el orden n con continuidad en el intervalo $[a,b]$ del ejercicio 312. Se representa, para cada $k = 1,\dots,n$, la k-ésima derivada de f por el símbolo $f^{(k)}$ y $f^{(0)}$ denota la función f. Se define, para cada $f \in C^n[a,b]$,

$$\|f\| = \sum_{k=0}^{n} \left\| f^{(k)} \right\|_\infty$$

donde $\left\| f^{(k)} \right\|_\infty$ es la norma uniforme en el espacio $C^k[a,b]$, para cada $k = 0,1,\dots,n$. Mostrar que $\|f\|$ es una norma en $C^n[a,b]$.

En los ejercicios 318 a 320, $C^n[a,b]$ y $\|f\| = \sum_{k=0}^{n} \left\| f^{(k)} \right\|_\infty$ son el espacio y la norma del ejercicio precedente.

318 Determinar si $f(x) = xe^{-x}$ pertenece a $C^1[0,3]$; en caso afirmativo calcular $\|f\|$

319 Determinar si $f(x) = \begin{cases} x^3 \ln x, \text{ si } x \neq 0 \\ 0, \text{ si } x = 0 \end{cases}$ pertenece a $C^1[0,3]$; en caso afirmativo calcular $\|f\|$.

320 Determinar si $f(x) = \begin{cases} e^{-1/x^2}, \text{ si } x \neq 0 \\ 0, \text{ si } x = 0 \end{cases}$ pertenece a $C^1[-1,1]$; en caso afirmativo calcular $\|f\|$.

321 Calcular $\|f\| = \text{máx}\{\|f\|_\infty, \|f'\|_\infty\}$, en $C^1[-\frac{\pi}{4}, \frac{\pi}{4}]$ si $f(x) = x \operatorname{sen} x$.

322 Calcular $\|f\| = \|f\|_\infty + \|f'\|_\infty$, en $C^1[-\frac{\pi}{4}, \frac{\pi}{4}]$ si $f(x) = x \operatorname{sen} x$.

323 Calcular $\|f\| = \text{máx}\{\|f\|_\infty, \|f'\|_\infty, \|f''\|_\infty\}$, en $C^2[-2,2]$ si $f(x) = e^{x^2}$.

324 Calcular $\|f\| = \|f\|_\infty + \|f'\|_\infty + \|f''\|_\infty$, en $C^2[-2,2]$ si $f(x) = e^{x^2}$.

325 Sea A una matriz y $\|A\| = \sqrt{\text{tra}(A^t A)}$ la norma canónica en los espacios de matrices (cfr. ejemplo 4.40, pág. 303). Probar que si $A \in \mathfrak{M}_{m\times n}$ y $B \in \mathfrak{M}_{n\times p}$ son cualquier par de matrices, entonces

$$\|AB\| \leq \|A\| \, \|B\| .$$

326 Sea $A = [a_{ij}]$ una matriz y la norma $\|A\|_1 = \sum |a_{ij}|$ donde la suma se toma sobre todos los subíndices i, j (cfr. ejemplo 4.44, pág. 306). Mostrar que si $A \in \mathfrak{M}_{m\times n}$ y $B \in \mathfrak{M}_{n\times p}$ son un par de matrices cualesquiera, entonces

$$\|AB\|_1 \leq \|A\|_1 \, \|B\|_1 .$$

327 Sea $A = [a_{ij}]$ es una matriz y $\|A\|_\infty = \text{máx}\, |a_{ij}|$, la norma cúbica definida en el ejemplo 4.43 (cfr. pág. 305). ¿Se cumple la relación de orden

$$\|AB\|_\infty \leq \|A\|_\infty \, \|B\|_\infty$$

para todo par de matrices $A \in \mathfrak{M}_{m\times n}$ y $B \in \mathfrak{M}_{n\times p}$?

328 Sea ℓ_1 el conjunto de todas las sucesiones $(a_n) \in \mathbb{R}^\infty$ tales que la serie $\sum_{n=1}^{\infty} |a_n|$ converge.

(a) Probar que ℓ_1 es un subespacio de $\mathbb{R}^\infty$.

(b) Se define, para cada $(a_n) \in \ell_1$,

$$\|(a_n)\|_1 = \sum_{n=1}^{\infty} |a_n|.$$

Mostrar que $\|(a_n)\|_1$ es una norma en ℓ_1.

En los ejercicios 329 a 338, ℓ_1 y $\|(a_n)\|_1 = \sum_{n=0}^{\infty} |a_n|$ son el espacio y la norma definidos en el ejercicio anterior.

329 Mostrar que $(a_n) \in \ell_1$ y calcular $\|(a_n)\|_1$ si $a_n = \frac{1}{n(n+1)}$.

330 Mostrar que $(a_n) \in \ell_1$ y calcular $\|(a_n)\|_1$ si $a_n = \frac{1}{16n^2-4}$.

331 Mostrar que $(a_n) \in \ell_1$ y calcular $\|(a_n)\|_1$ si $a_n = \frac{1}{n!}$.

332 Mostrar que $(a_n) \in \ell_1$ y calcular $\|(a_n)\|_1$ si $a_n = \frac{4^{n+1}}{(n-1)!}$.

333 Mostrar que $(a_n) \in \ell_1$ y calcular $\|(a_n)\|_1$ si $a_n = \frac{1}{2^{n-1}}$.

334 Mostrar que $(a_n) \in \ell_1$ y calcular $\|(a_n)\|_1$ si $a_n = (-3)^{n-4}5^{2-n}$.

335 Mostrar que $(a_n) \in \ell_1$ y calcular $\|(a_n)\|_1$ si $a_n = \frac{2^n}{n!}$.

336 Mostrar que $(a_n) \in \ell_1$ y calcular $\|(a_n)\|_1$ si $a_n = \frac{1}{4^{2n}}$.

337 Mostrar que $(a_n) \in \ell_1$ y calcular $\|(a_n)\|_1$ si $a_n = \left(\frac{1}{2^n}\right)\frac{1}{n}$.

338 Mostrar que $(a_n) \in \ell_1$ y calcular $\|(a_n)\|_1$ si $a_n = \left(\frac{1}{3^{n-1}}\right)\frac{1}{n}$.

339 Sea $p \geq 1$ un número real dado y ℓ_p el conjunto de sucesiones $(a_n) \in \mathbb{R}^\infty$ tales que

$$\sum_{n=1}^{\infty} |a_n|^p$$

converge. Se define, para cada $(a_n) \in \ell_p$,

$$\|(a_n)\|_p = \left(\sum_{n=1}^{\infty} |a_n|^p\right)^{1/p}.$$

(a) Mostrar que ℓ_p es un subespacio de $\mathbb{R}^\infty$.

(b) Demostrar que $\|(a_n)\|_p$ es una norma en ℓ_p.

(c) Probar que si $p, p^* > 1$ son índices conjugados; i.e., $\frac{1}{p} + \frac{1}{p^*} = 1$, y $(a_n) \in \ell_p$, $(b_n) \in \ell_{p^*}$, entonces $(a_n b_n) \in \ell_1$ y además

$$\sum_{n=1}^{\infty} |a_n b_n| \le \|(a_n)\|_p \|(b_n)\|_{p^*}$$

es decir,

$$\|(a_n b_n)\|_1 \le \|(a_n)\|_p \|(b_n)\|_{p^*}$$

(Sugerencia: Utilizar las desigualdades de Hölder (4.50) y Minkowski (4.51), pág. 330.)

En los ejercicios 340 a 345, ℓ_p y $\|\cdot\|_p$ son el espacio y la norma definidos en el ejercicio anterior. Determinar si la sucesión (a_n) pertenece a ℓ_p y, en caso afirmativo, calcular $\|(a_n)\|_p$ si:

340 $a_n = \dfrac{1}{2^n}$, $p = 3$.

341 $a_n = \dfrac{1}{\sqrt{n}}$, $p = 2$.

342 $a_n = (-1)^n \sqrt{\dfrac{1}{n!}}$, $p = 2$.

343 $a_n = r^n$, $p > 1$ cualquier número real mayor que 1 y $-1 < r < 1$.

344 $a_n = \sqrt[5]{\dfrac{n+1}{3^n}}$, $p = 5$.

345 $a_n = \sqrt[3]{\dfrac{1}{2^n n}}$, $p = 3$.

En los ejercicios 346 a 355, $\|f\|_p = \left(\int_a^b |f(x)|^p\, dx\right)^{1/p}$ es la norma definida en el ejercicio resuelto 35 para un número $p > 1$. Calcular $\|f\|_p$ para la función f y el número p en el espacio indicado.

346 $f(x) = x^{-1}$, $p > 1$ en $C[1,2]$.

347 $f(x) = x$, $p > 1$ en $C[0,1]$.

348 $f(x) = xe^x$, $p = 2$ en $C[0,1]$.

349 $f(x) = x\cos x$, $p = 2$ en $C[0,2\pi]$.

350 $f(x) = x\,\mathrm{sen}^{1/3} x$, $p = 3$ en $C[0,2\pi]$.

351 $f(x) = x\ln^{2/3} x$, $p = 3/2$ en $C[1,e]$.

352 $f(x) = \cos^{1/3} x \operatorname{sen} x$, $p = 3$ en $C[0,\pi/2]$.

353 $f(x) = e^x \operatorname{sen}^{1/4} x$, $p = 4$ en $C[0,2\pi]$.

354 $f(x) = \tan x$, $p = 2$ en $C[0,\pi/4]$.

355 $f(x) = \tan^{-1/4} x$, $p = 2$ en $C[0,\pi/4]$.

356 Sean $p \geq 1$ y a un par de números reales dados y $\mathscr{L}_p[a,\infty)$ el conjunto de funciones reales continuas en $[a,\infty)$ tales que $\int_a^\infty |f(x)|^p\,dx$ es convergente. Se define, para cada

$$f \in \mathscr{L}_p[a,\infty), \|f\|_p = \left(\int_a^\infty |f(x)|^p\,dx\right)^{1/p}.$$

(a) Mostrar que $\mathscr{L}_p[a,\infty)$ es un subespacio de $C[a,\infty)$ y, por tanto, un espacio vectorial con la suma de funciones y multiplicación de un escalar con una función usuales.

(b) Probar que $\|f\|_p$ es una norma en $\mathscr{L}_p[a,\infty)$.

(c) Si $p, p^* > 1$ son índices conjugados ($\frac{1}{p} + \frac{1}{p^*} = 1$), $f \in \mathscr{L}_p[a,\infty)$ y $g \in \mathscr{L}_{p^*}[a,\infty)$, demostrar que $fg \in \mathscr{L}_1[a,\infty)$ (cfr. ejercicio resuelto 37) y que

$$\int_a^\infty |fg| \leq \|f\|_p \|g\|_{p^*}$$

En los ejercicios 357 a 363, $\mathscr{L}_p[a,\infty)$ y $\|f\|_p = \left(\int_a^\infty |f(x)|^p\,dx\right)^{1/p}$ son el espacio y la norma definidos en el ejercicio precedente. Determinar si la función f pertenece al espacio $\mathscr{L}_p[a,\infty)$ y en caso afirmativo calcular $\|f\|_p$ si:

357 $f(x) = \dfrac{1}{x}$, $p = 1$ y $a = 1$.

358 $f(x) = \dfrac{1}{(x-1)^2}$, $p = 1$ y $a = 2$.

359 $f(x) = \dfrac{1}{x}$, $p > 1$ y $a > 1$.

360 $f(x) = \dfrac{\ln x}{x}$, $p = 1$ y $a = 1$.

361 $f(x) = \dfrac{(\ln x)^{1/2}}{x}$, $p = 2$ y $a = 1$.

362 $f(x) = xe^{-x/p}$, $p = 1, 2, \ldots$ y $a = 0$.

363 $f(x) = \dfrac{(\ln x)^{-1}}{\sqrt{x}}$, $p = 2$, $a = 2$.

Una seminorma en un espacio vectorial $\mathbf{E}$ es una función $\mathscr{N} : \mathbf{E} \to \mathbb{R}$ tal que:

- $\mathscr{N}(\vec{u}) \geq 0\ \forall \vec{u} \in \mathbf{E}$.
- $\mathscr{N}(\lambda\vec{u}) = |\lambda|\,\vec{u}$ para todo $\lambda \in \mathbb{R}$ y para todo $\vec{u} \in \mathbf{E}$.

- $\mathscr{N}(\vec{u}+\vec{v}) \leq \mathscr{N}(\vec{u})+\mathscr{N}(\vec{v})$.
- Existe $\vec{u} \in \mathbf{E}$, con $\vec{u} \neq \vec{0}_{\mathbf{E}}$, tal que $\mathscr{N}(\vec{u}) = 0$.

Probar, en los ejercicios 364 a 366, que el conjunto **E** es un espacio vectorial (con las operaciones usuales) y que $\mathscr{N}(\vec{u})$ es una seminorma en él.

364 (**Seminorma de Taylor**). Sean n un entero positivo, J un intervalo abierto que contiene a un punto dado a, **E** el conjunto de funciones $f : J \to \mathbb{R}$ tales que f tiene derivada hasta el orden n en el punto a y

$$\mathscr{N}(f) = \sum_{k=0}^{n} \left| f^{(k)}(a) \right|.$$

donde $f^{(0)}(a) = f(a)$ y $f^{(k)}(a)$ es la k-ésima derivada de f en a, $k = 1, \ldots, n$.

365 (**Seminorma de interpolación**). Sean n un entero positivo fijo, $\mathbf{E} = \mathscr{F}([a,b])$ el conjunto de funciones $f : [a,b] \to \mathbb{R}$, $x_0, x_1, \ldots, x_n \in [a,b]$ puntos fijos distintos entre sí y

$$\mathscr{N}(f) = \sum_{k=0}^{n} |f(x_k)|.$$

366 Sean n un entero positivo, J un intervalo abierto que contiene a un punto dado a, **E** el conjunto de funciones $f : J \to \mathbb{R}$ tales que f tiene derivada hasta el orden n en el punto a y

$$\mathscr{N}(f) = \sum_{k=0}^{n} \frac{1}{k!} \left| f^{(k)}(a) \right|$$

donde $f^{(0)}(a) = f(a)$ y $f^{(k)}(a)$ es la k-ésima derivada de f en a, $k = 1, \ldots, n$.

367 Demostrar que la seminorma del ejercicio 366 satisface la relación de orden

$$\mathscr{N}(fg) \leq \mathscr{N}(f)\mathscr{N}(g)$$

para cualquier par de funciones $f, g \in \mathbf{E}$.

368 ¿Satisfacen las seminormas de Taylor e interpolación de los ejercicios 364 y 365 la relación de orden $\mathscr{N}(fg) \leq \mathscr{N}(f)\mathscr{N}(g)$?

369 Sea **E** un espacio vectorial y $\mathscr{N}$ una seminorma en él.

(a) Probar el conjunto $S_{\mathscr{N}}$ de todos los $f \in \mathbf{E}$ tales que $\mathscr{N}(f) = 0$, es un subespacio de **E**.

(b) Sea $C^1[a,b]$ el espacio de funciones con primera derivada continua en el intervalo $[a,b]$. Probar que

$$\mathscr{N}(f) = \max_{a \leq x \leq b} |f'(x)|$$

es una seminorma en $C^1[a,b]$. Determinar $S_{\mathscr{N}}$ para este caso.

370 Sean **E** un espacio vectorial, $\mathscr{N}$ una seminorma y $\mathscr{N}_1$ una norma en él. Probar que

$$\|\vec{u}\| = \mathscr{N}(\vec{u}) + \mathscr{N}_1(\vec{u})$$

es una norma en **E**.

Sea $C[a,b]$ el espacio de funciones continuas en $[a,b]$ dotado de la norma uniforme $\|f\|_\infty = \max_{a\le x\le b} |f(x)|$. En los ejercicios 371 a 374, bosquejar $B[f,\varepsilon]$, la bola de centro f y radio ε respecto a esta norma para cada f dada.

371 $f(x) = 1$, $a = 0$ y $b = 2$.

372 $f(x) = x^2$, $a = -1$ y $b = 1$.

373 $f(x) = e^{-x}$, $a = -1$ y $b = 2$.

374 $f(x) = \cos x$, $a = -\pi$ y $b = \pi$.

En los ejercicios 375 a 384, se define una función $(x_1,\ldots,x_n) \mapsto \|(x_1,\ldots,x_n)\|$ en el espacio $\mathbb{R}^n$ indicado. Utilizar los teoremas 4.17 y 4.19 (cfr. págs. 318 y 320) para:

(a) Probar que $\|(x_1,\ldots,x_n)\|$ es una norma y que proviene de un producto interior.

(b) Hallar el producto interior del cual proviene la norma.

375 $\|(x_1,x_2)\| = 2x_1^2 - 2x_2x_1 + x_2^2$ en $\mathbb{R}^2$.

376 $\|(x_1,x_2)\| = x_1^2 - 2x_2x_1 + 2x_2^2$ en $\mathbb{R}^2$.

377 $\|(x_1,x_2)\| = 5x_1^2 + 6x_1x_2 + 2x_2^2$ en $\mathbb{R}^2$.

378 $\|(x_1,x_2)\| = x_1^2 - 6x_1x_2 + 10x_2^2$ en $\mathbb{R}^2$.

379 $\|(x_1,x_2,x_3)\| = 5x_1^2 - 2x_1x_2 - 4x_1x_3 + 2x_2^2 + x_3^2$ en $\mathbb{R}^3$.

380 $\|(x_1,x_2,x_3)\| = 2x_1^2 - 2x_1x_3 + x_2^2 + 2x_2x_3 + 2x_3^2$ en $\mathbb{R}^3$.

381 $\|(x_1,x_2,x_3)\| = x_1^2 + x_2^2 + 2x_2x_3 + 2x_3^2$ en $\mathbb{R}^3$.

382 $\|(x_1,x_2,x_3)\| = 2x_1^2 - 2x_1x_2 - 2x_1x_3 + 5x_2^2 + x_3^2$ en $\mathbb{R}^3$.

383 $\|(x_1,x_2,x_3,x_4)\| = x_1^2 + 2x_2^2 + 2x_2x_3 - 2x_2x_4 + x_3^2 + 2x_4^2$ en $\mathbb{R}^4$.

384 $\|(x_1,x_2,x_3,x_4,x_5)\| = x_1^2 - 2x_1x_5 + x_2^2 + 2x_3^2 - 2x_3x_5 + x_4^2 + 2x_5^2$ en $\mathbb{R}^5$.

En los ejercicios 385 a 398, determinar si la norma en el espacio indicado proviene o no de un producto interior; en caso afirmativo, hallar el producto escalar del cual proviene la norma (cfr. teoremas 4.17 y 4.19).

385 $\|(x_1,\ldots,x_n)\|_1 = \sum_{i=1}^n |x_i|$ en $\mathbb{R}^n$.

386 $\|(a_{ij})\|_1 = \sum |a_{ij}|$ en $\mathfrak{M}_{2\times 3}$.

387 $\|(a_{ij})\|_\infty = \text{máx}\,|a_{ij}|$ en $\mathfrak{M}_{3\times 2}$.

388 $\|f\|_\infty = \max_{0\le x\le 1} |f(x)|$ en $C[0,1]$.

389 $\|f\|_1 = \int_0^1 |f(x)|\,dx$ en $C[0,1]$.

390 $\|f\|_2 = \left(\int_0^1 [f(x)]^2 dx\right)^{1/2}$ en $C[0,1]$.

391 $\|f\|_3 = \left(\int_0^1 [f(x)]^3 dx\right)^{1/3}$ en $C[0,1]$.

392 $\|(a_n)\|_1 = \sum_{n=0}^{\infty} |a_n|$ en ℓ_1.

393 $\|(a_n)\|_2 = \left(\sum_{n=0}^{\infty} |a_n|^2\right)^{1/2}$ en ℓ_2.

394 $\|(a_n)\|_3 = \left(\sum_{n=0}^{\infty} |a_n|^3\right)^{1/3}$ en ℓ_3.

395 $\|(a_n)\|_p = \left(\sum_{n=0}^{\infty} |a_n|^p\right)^{1/p}$ en ℓ_p, con $p \neq 2$.

396 $\|f\| = \left(\int_0^\infty [f(x)]^2 e^{-x} dx\right)^{1/2}$ en el espacio vectorial de las funciones continuas $f:[0,\infty)\to\mathbb{R}$ tales que $\int_0^\infty [f(x)]^2 e^{-x} dx$ converge.

397 $\|f\| = \max(\|f\|_\infty, \|f'\|_\infty)$ en $C^1[0,1]$.

398 $\|f\| = \|f\|_\infty + \|f'\|_\infty$ en $C^1[0,1]$.

399 Demostrar, sin utilizar el teorema 4.22 (cfr. pág. 327), que las normas

$$\|(x_1,\ldots,x_n)\|_1 = \sum_{i=1}^{n} |x_i| \quad \text{y} \quad \|(x_1,\ldots,x_n)\|_\infty = \max_{1\le i\le n} |x_i|$$

del espacio $\mathbb{R}^n$, son equivalentes.

400 Demostrar, sin utilizar el teorema 4.22, que las normas

$$\|(x_1,\ldots,x_n)\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \quad \text{y} \quad \|(x_1,\ldots,x_n)\|_\infty = \max_{1\le i\le n} |x_i|$$

(del espacio $\mathbb{R}^n$) son equivalentes.

401 Encontrar (sin utilizar el teorema 4.22) $\alpha > 0$ y $\beta > 0$ tales que para todo $\vec{u} = (x_1,\ldots,x_n) \in \mathbb{R}^n$ se cumpla $\alpha\|\vec{u}\| \le \|\vec{u}\|_1 \le \beta\|u\|$ donde $\|\vec{u}\| = \left(\sum_{i=1} x_i^2\right)^{1/2}$ y $\|\vec{u}\|_1 = \sum_{i=1} |x_i|$.

402 Sean $C[0,1]$ el espacio de funciones continuas en el intervalo $[0,1]$,

$$\|f\|_2 = \left(\int_0^1 |f(x)|^2 dx\right)^{1/2} \quad \text{y} \quad \|f\|_\infty = \max_{a\le x\le b} |f(x)|.$$

(a) Encontrar $\beta > 0$ tal que

$$\|f\|_2 \le \beta \|f\|_\infty$$

para todo $f \in C[0,1]$.

(b) Demostrar que no puede existir $\alpha > 0$ tal que

$$\alpha \|f\|_\infty \le \|f\|_2$$

para todo $f \in C[0,1]$.

(c) ¿Son equivalentes $\|f\|_2$ y $\|f\|_\infty$?

403 Sean las normas $\|f\|_1 = \int_a^b |f(x)|\,dx$ y $\|f\|_2 = \left(\int_a^b [f(x)]^2 dx\right)^{1/2}$ en el espacio $C[a,b]$. Probar que

$$\|f\|_1 \le \sqrt{b-a}\,\|f\|_2$$

para toda función $f \in C[a,b]$.

404 Probar que las normas

$$\|f\|_1 = \int_0^1 |f(x)|\,dx \quad \text{y} \quad \|f\|_2 = \left(\int_0^1 [f(x)]^2 dx\right)^{1/2}$$

en el espacio $C[0,1]$ no son equivalentes.

Sean $\mathbf{E}$ un espacio vectorial y $\vec{a},\vec{b} \in \mathbf{E}$. Se define el segmento con punto inicial $\vec{a}$ y punto final $\vec{b}$ como el conjunto $[\vec{a},\vec{b}] = \left\{\vec{u} \in \mathbf{E} \,|\, \vec{u} = \vec{a} + t(\vec{b}-\vec{a}),\, 0 \le t \le 1\right\}$.

405 Un conjunto $M \subset \mathbf{E}$ es convexo si el segmento $[\vec{a},\vec{b}] \in M$ para todo par $\vec{a},\vec{b} \in M$.

(a) Probar que todo subespacio S de $\mathbf{E}$ es un conjunto convexo.

(b) Sean $r > 0$ y $\vec{u}_0 \in \mathbf{E}$, demostrar que la bola cerrada $B[\vec{u}_0, r]$ es un conjunto convexo.

406 Sean $\alpha, a_i \in \mathbb{R}$, $i = 1,\dots,n$, números reales dados. Se llama semiplano al conjunto de puntos $(x_1,\dots,x_n) \in \mathbb{R}^n$ tales que

$$a_1x_1 + \cdots + a_nx_n \ge \alpha.$$

(a) En referencia al ejercicio 405, probar que todo semiplano es convexo.

(b) Demostrar que el complemento de un semiplano (al que también se le llama semiplano) es un conjunto convexo.

407 Probar que toda intersección finita de conjuntos convexos en un espacio vectorial es también un conjunto convexo (cfr. ejercicio 405).

408 Sean S un subespacio de dimensión finita de un espacio normado $\mathbf{E}$ y $\vec{u} \in \mathbf{E} - S$. Probar que el conjunto S^* de las aproximaciones óptimas de $\vec{u}$ en S es un conjunto convexo (cfr. ejercicio 405).

409 Sea la función continua, en el intervalo $[0,1]$, $f(x) = x\cos x + 4$. Calcular los polinomios de Bernstein B_1, B_2, B_3, B_4 y B_5 (cfr. pág. 339) para esta función.

410 Sea $f(x) = e^x$. $f \in C[0,1]$; calcular los polinomios de Bernstein B_i, $i = 1,2,3,4$.

411 Sea $f(x) = x^3$; $f \in C[0,1]$; calcular los polinomios de Bernstein B_i, $i = 1,2,3,\ldots$ para f. Mostrar directamente que para este caso

$$\lim_{n\to\infty} \|f - B_n\|_\infty = 0.$$

412 Sea $f(x) = \operatorname{sen} x$, $0 \le x \le 1$. Calcular los polinomios de Bernstein B_i, $i = 1,2,3$, para la función f (cfr. pág. 339).

413 Sea $f(x) = \cos x$, $0 \le x \le 1$. Calcular los polinomios de Bernstein B_i, $i = 1,2,3$, para la función f (cfr. pág. 339).

414 Sean n un entero positivo, J un intervalo abierto que contiene un punto dado a, $\mathbf{E}$ el espacio de las funciones que tienen derivada hasta el orden n en el punto a y

$$\mathcal{N}(f) = \sum_{k=0}^{n} \left| f^{(k)}(a) \right|$$

la seminorma de Taylor (cfr. ejercicio 364). Sean $f \in \mathbf{E}$ y el polinomo de grado a lo más n

$$p_n^*(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

(a) Probar que $p_n^*(x)$ es aproximación óptima de f en el subespacio $\mathbf{P}_n$ de $\mathbf{E}$ respecto a la seminorma de Taylor.

(b) Mostrar que si $p_1^*(x)$ también aproximación óptima de f en $\mathbf{P}_n$, entonces $p^* = p_1^*$.

En los ejercicios 415 a 417, $S = \operatorname{gn}((-1,1))$ y $\vec{u} = (1,1)$ en el espacio $\mathbb{R}^2$.

415 Hallar la aproximación óptima $\vec{p}^*$ de $\vec{u}$ en S respecto a la norma euclidiana $\|(x,y)\| = \sqrt{x^2+y^2}$.

416 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma cúbica $\|(x,y)\|_\infty = \text{máx}\,\{|x|,|y|\}$.

417 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma $\|(x,y)\|_1 = |x| + |y|$.

En los ejercicios 418 a 420, $S = \operatorname{gn}((2,1))$ y $\vec{u} = (-1,1)$ en el espacio $\mathbb{R}^2$.

418 Hallar la aproximación óptima $\vec{p}^*$ de $\vec{u}$ en S respecto a la norma euclidiana $\|(x,y)\| = \sqrt{x^2+y^2}$.

419 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma cúbica $\|(x,y)\|_\infty = \text{máx}\,\{|x|,|y|\}$.

420 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma $\|(x,y)\|_1 = |x| + |y|$.

En los ejercicios 421 a 423, $S = \operatorname{gn}((-2,1))$ y $\vec{u} = (3,1)$ en el espacio $\mathbb{R}^2$.

421 Hallar la aproximación óptima $\vec{p}^*$ de $\vec{u}$ en S respecto a la norma euclidiana $\|(x,y)\| = \sqrt{x^2+y^2}$.

422 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma cúbica $\|(x,y)\|_\infty = \text{máx}\,\{|x|,|y|\}$.

423 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma $\|(x,y)\|_1 = |x| + |y|$.

En los ejercicios 424 a 426, $S = \text{gn}((-1,1,2))$ y $\vec{u} = (1,1,1)$ en el espacio $\mathbb{R}^3$.

424 Hallar la aproximación óptima $\vec{p}^*$ de $\vec{u}$ en S respecto a la norma euclidiana $\|(x,y,z)\| = \sqrt{x^2+y^2+z^2}$.

425 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma cúbica

$$\|(x,y,z)\|_\infty = \text{máx}\{|x|,|y|,|z|\}.$$

426 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma $\|(x,y,z)\|_1 = |x| + |y| + |z|$.

En los ejercicios 427 a 429, $S = \text{gn}(1,-1,3)$ y $\vec{u} = (1,0,-1)$ en el espacio $\mathbb{R}^3$.

427 Hallar la aproximación óptima $\vec{p}^*$ de $\vec{u}$ en S respecto a la norma euclidiana $\|(x,y,z)\| = \sqrt{x^2+y^2+z^2}$.

428 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma cúbica

$$\|(x,y,z)\|_\infty = \text{máx}\{|x|,|y|,|z|\}.$$

429 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma $\|(x,y,z)\|_1 = |x| + |y| + |z|$.

En los ejercicios 430 a 432, $S = \text{gn}((0,-1,-2))$ y $\vec{u} = (1,-1,1)$ en el espacio $\mathbb{R}^3$.

430 Hallar la aproximación óptima $\vec{p}^*$ de $\vec{u}$ en S respecto a la norma euclidiana $\|(x,y,z)\| = \sqrt{x^2+y^2+z^2}$.

431 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma cúbica

$$\|(x,y,z)\|_\infty = \text{máx}\{|x|,|y|,|z|\}.$$

432 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma $\|(x,y,z)\|_1 = |x| + |y| + |z|$.

En los ejercicios 433 a 434, S es el plano $y - z = 0$ y $\vec{u} = (0,1,0)$ en el espacio $\mathbb{R}^3$.

433 Hallar la aproximación óptima $\vec{p}^*$ de $\vec{u}$ en S respecto a la norma euclidiana $\|(x,y,z)\| = \sqrt{x^2+y^2+z^2}$.

434 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma $\|(x,y,z)\|_1 = |x| + |y| + |z|$.

En los ejercicios 435 a 436, S es el plano $x + y - z = 0$ y $\vec{u} = (0,1,0)$ en el espacio $\mathbb{R}^3$.

435 Hallar la aproximación óptima $\vec{p}^*$ de $\vec{u}$ en S respecto a la norma euclidiana $\|(x,y,z)\| = \sqrt{x^2+y^2+z^2}$.

436 Hallar todas las aproximaciones óptimas de $\vec{u}$ en S respecto a la norma $\|(x,y,z)\|_1 = |x| + |y| + |z|$.

En los ejercicios 437 a 439, $S = \text{gn}(1)$ y $f(x) = x$ en el espacio $C[0,1]$.

437 Hallar la aproximación óptima, p^*, de f en S respecto a la norma $\|g\| = \left(\int_0^1 (g(x))^2 dx\right)^{1/2}$.

438 Hallar todas las aproximaciones óptimas de f en S respecto a la norma cúbica $\|g\|_\infty = \text{máx}_{0\leq x\leq 1}|g(x)|$.

439 Hallar todas las aproximaciones óptimas de f en S respecto a la norma $\|g\|_1 = \int_0^1 |g(x)|\,dx$.

En los ejercicios 440 a 442, $S = \text{gn}(1)$ y $f(x) = \ln x$ en el espacio $C[1,e]$.

440 Hallar la aproximación óptima, p^*, de f en S respecto a la norma $\|g\| = \left(\int_1^e (g(x))^2 dx\right)^{1/2}$.

441 Encontrar todas las aproximaciones óptimas de f en S respecto a la norma uniforme $\|g\|_\infty$.

442 Encontrar todas las aproximaciones óptimas de f en S respecto a la norma $\|g\|_1 = \int_1^e |g(x)|\,dx$.

En los ejercicios 443 a 445, $S = \text{gn}(1)$ y $f(x) = \cos x$ en el espacio $C[0,\pi/2]$.

443 Hallar la aproximación óptima, p^*, de f en S respecto a la norma $\|g\| = \left(\int_0^{\pi/2} (g(x))^2 dx\right)^{1/2}$.

444 Encontrar todas las aproximaciones óptimas de f en S respecto a la norma uniforme $\|g\|_\infty$.

445 Encontrar todas las aproximaciones óptimas de f en S respecto a la norma $\|g\|_1 = \int_0^{\pi/2} |g(x)|\,dx$.

En los ejercicios 446 a 448, $S = \text{gn}(1)$ y $f(x) = \operatorname{sen} x$ en el espacio $C[0,\pi/2]$.

446 Hallar la aproximación óptima, p^*, de f en S respecto a la norma $\|g\| = \left(\int_0^{\pi/2} (g(x))^2 dx\right)^{1/2}$.

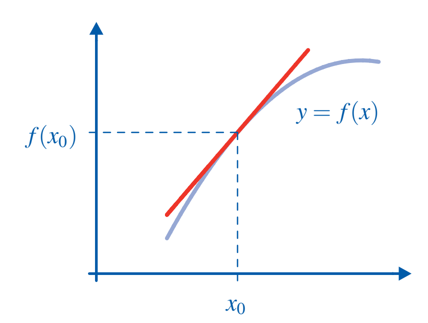
447 Encontrar todas las aproximaciones óptimas de f en S respecto a la norma uniforme $\|g\|_\infty$.

448 Encontrar todas las aproximaciones óptimas de f en S respecto a la norma $\|g\|_1 = \int_0^{\pi/2} |g(x)|\,dx$.

5 Transformaciones lineales, valores y vectores propios

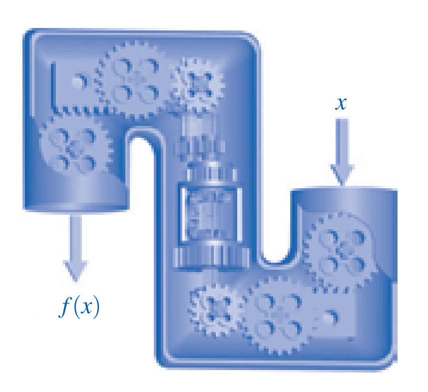
En las primeras secciones de este capítulo estudiaremos cierto tipo de funciones entre espacios vectoriales: las transformaciones lineales. Estas funciones son relativamente muy simples de tratar ya que exhiben un comportamiento que preserva la estructura de las operaciones de espacio vectorial. A pesar de su sencillez, las transformaciones lineales son muy importantes tanto en matemáticas como en física, ingeniería y ciencias sociales. Podríamos afirmar, *grosso modo*, que independientemente de la gran variedad de sus aplicaciones, mucho del éxito que tienen las funciones lineales entre espacios vectoriales radica en que con frecuencia pueden transformar un problema complejo en uno más simple. En las subsecuentes secciones trataremos el tema no menos importante, y estrechamente relacionado con las transformaciones lineales, de valores y vectores propios. Como antes, la última sección está dedicada a ejercicios resueltos y a ejercicios propuestos al lector.

5.1 Transformaciones lineales



Las funciones más sencillas (después de las constantes) de una variable con valores reales son las funciones de la forma $f(x) = kx$; cuya gráfica, para un valor fijo k, es una línea recta con pendiente k que pasa por el origen. La forma simple que tienen estas funciones las hace sumamente importantes para estudiar el comportamiento de funciones más complicadas. Por ejemplo, una función que es derivable en un punto x_0 se puede aproximar localmente por medio de la línea recta tangente a la gráfica de la función en el punto $(x_0, f(x_0))$; y esta línea recta no es más que una traslación afín de la recta $y = f'(x_0)x$. Una gran variedad de fenómenos se pueden modelar a través de soluciones de cierto tipo de ecuaciones que exhiben un comportamiento lineal, en el sentido de que la suma de dos soluciones y el producto de un escalar por una solución también son soluciones; dichos fenómenos y sus respectivos modelos son llamados, por antonomasia, lineales también. Con base en la característica de linealidad de estos fenómenos es posible, en general, determinar el comportamiento de los mismos en una forma relativamente sencilla. Así, como las funciones lineales de una variable sirven para aproximar funciones más complicadas, los modelos lineales se pueden utilizar para aproximar fenómenos más complejos. Por sí solos los fenómenos lineales son sumamente interesantes y cubren una gran variedad de importantes aplicaciones. Las funciones lineales de una variable tienen una inmediata extensión a funciones de varias variables y, más aún, a funciones entre espacios vectoriales. En esta sección estudiaremos en un contexto general este tipo de funciones que llamaremos transformaciones lineales.

5.1.1 Definición, ejemplos y propiedades



Recordemos que si A y B son un par de conjuntos no vacíos, la notación $f : A \to B$ significa que f es una función con dominio A y valores en B o, en forma más compacta, *f es una función de A en B* (cfr. definición 3.7, pág. 133); que en la notación $y = f(x)$ a y se le llama el valor de la función f en x o la imagen de x bajo la función f; a x se le dice variable independiente (o argumento de la función) y a y variable dependiente; y que al conjunto B se le dice contradominio de la función. Una manera muy útil de interpretar una función f es como un conjunto de procedimientos, una máquina, que transforma cada elemento x (materia prima) de A en un producto $y = f(x)$ elemento de B. El resultado de esta transformación, el elemento $y = f(x)$, depende de la materia prima x que se introduzca cada vez a la máquina, y esta transformación está perfectamente determinada por esta máquina o conjunto de procedimientos f (la función o transformación). En este sentido interpretaremos, de aquí en adelante, el significado de una función entre espacios vectoriales: *como una transformación de un espacio en otro*.

Definición 5.1 ***(Transformación lineal)*** *Sean* $\mathbf{E}$ *y* $\mathbf{F}$ *dos espacios vectoriales y* $T : \mathbf{E} \to \mathbf{F}$ *una función. Diremos que* T *es una* ***transformación lineal*** *de* $\mathbf{E}$ *en* $\mathbf{F}$ *si:*[1]

1. $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y}), \forall \vec{x}, \vec{y} \in \mathbf{E}$.
2. $T(\alpha\vec{x}) = \alpha T(\vec{x}), \forall \vec{x} \in \mathbf{E}, \forall \alpha \in \mathbb{R}$.

► **Ejemplo 5.1** Sea $T : \mathbb{R} \to \mathbb{R}$ la función definida por $T(x) = 3x$, entonces para $x_1, x_2 \in \mathbb{R}$ y para todo escalar α,

$$\begin{aligned} T(x_1 + x_2) &= 3(x_1 + x_2) \\ &= 3x_1 + 3x_2 \\ &= T(x_1) + T(x_2) \end{aligned}$$

y

$$\begin{aligned} T(\alpha x) &= 3(\alpha x) \\ &= \alpha(3x) \\ &= \alpha T(x). \end{aligned}$$

Por tanto, T es una transformación lineal.◄

► **Ejemplo 5.2** De manera análoga al ejemplo precedente, toda línea recta que pasa por el origen es una transformación lineal de $\mathbb{R}$ en $\mathbb{R}$. En efecto, una línea recta que pasa por el origen es la gráfica de una función de la forma $y = T(x) = ax$, donde a es una constante. Entonces, si $x_1, x_2 \in \mathbb{R}$ y α es un escalar, se tiene

[1] Se le dice también aplicación lineal u homomorfismo.

$$\begin{aligned}T(x_1+x_2) &= a(x_1+x_2)\\ &= ax_1+ax_2\\ &= T(x_1)+T(x_2)\end{aligned}$$

y

$$\begin{aligned}T(\alpha x_1) &= a(\alpha x_1)\\ &= \alpha(ax_1)\\ &= \alpha T(x_1).\end{aligned}$$ ◀

Vimos en el ejemplo anterior que toda línea recta que pasa por el origen es una transformación lineal de $\mathbb{R}$ en $\mathbb{R}$; de hecho, estas líneas rectas son las únicas transformaciones lineales que existen de $\mathbb{R}$ en $\mathbb{R}$, como hacemos patente en el siguiente ejemplo.

▶ **Ejemplo 5.3** (Transformaciones lineales de $\mathbb{R}$ en $\mathbb{R}$) Sea $T:\mathbb{R}\to\mathbb{R}$ una transformación lineal. Sea $x\in\mathbb{R}$, entonces $T(x)=T(x\cdot 1)=xT(1)$. Sea $a=T(1)$, entonces $T(x)=ax\ \forall x\in\mathbb{R}$. Es decir, T es una línea recta que pasa por el origen.◀

▶ **Ejemplo 5.4** (La transformación derivación) Si $\mathbf{E}=C^1[0,1]$ (cfr. el ejemplo 3.24, pág. 142) y $\mathbf{F}=C[0,1]$, sea $T:\mathbf{E}\to\mathbf{F}$, definida por $T(f)=f'$. Así, por ejemplo, si $f(x)=x^2$ y $g(x)=2x$, $T(f)=g$. T es una transformación lineal. En efecto:

1. $T(f+g)=(f+g)'=f'+g'=T(f)+T(g)\ \forall f,g\in\mathbf{E}$.
2. $T(\alpha f)=(\alpha f)'=\alpha f'\ \forall\alpha\in\mathbb{R}, \forall f\in\mathbf{E}$.◀

▶ **Ejemplo 5.5** Sea $T:\mathbb{R}^3\to\mathbb{R}^2$ definida como $T(x_1,x_2,x_3)=(x_1-2x_2,x_2+3x_3)$. Mostrar que T es lineal.◀

DEMOSTRACIÓN ■ 1. Si $\vec{a}=(x_1,x_2,x_3)$ y $\vec{b}=(y_1,y_2,y_3)$,

$$\begin{aligned}T(\vec{a}+\vec{b}) &= T(x_1+y_1,x_2+y_2,x_3+y_3)\\ &= (x_1+y_1-2(x_2+y_2),x_2+y_2+3(x_3+y_3)).\\ T(\vec{a})+T(\vec{b}) &= T(x_1,x_2,x_3)+T(y_1,y_2,y_3)\\ &= (x_1-2x_2,x_2+3x_3)+(y_1-2y_2,y_2+3y_3)\\ &= (x_1+y_1-2(x_2+y_2),x_2+y_2+3(x_3+y_3))\\ &= T(\vec{a}+\vec{b}).\end{aligned}$$

2. Si $\vec{a}=(x_1,x_2,x_3)$ y $\alpha\in\mathbb{R}$,

$$\begin{aligned}T(\alpha\vec{a}) &= T(\alpha(x_1,x_2,x_3))\\ &= T(\alpha x_1,\alpha x_2,\alpha x_3)\\ &= T(\alpha x_1-2\alpha x_2,\alpha x_2+3\alpha x_3)\\ &= \alpha(x_1-2x_2,x_2+3x_3)\\ &= \alpha T(x_1,x_2,x_3)\\ &= \alpha T(\vec{a}).\end{aligned}$$ ■

▶ **Ejemplo 5.6** Sea $T:\mathbb{R}^2\to\mathbb{R}^2$, definida por $T(x,y)=xy$, T no es lineal:

$$\begin{aligned} T(3(1,2)) &= T(3,6)\\ &= 18.\\ 3T(1,2) &= 3\cdot 2\\ &= 6.\end{aligned}$$

Por lo que $T(3(1,2))\neq 3T(1,2)$.◀

▶ **Ejemplo 5.7** (Transformación integración) Sea $T:C[0,1]\to\mathbb{R}$ definida por:

$$T(f)=\int_0^1 f(x)dx.$$

Así, por ejemplo, si $f(x)=x^2$, $T(f)=1/3$. T es lineal porque:

1. Si $f,g\in C[0,1]$,

$$\begin{aligned} T(f+g) &= \int_0^1 (f+g)(x)dx\\ &= \int_0^1 (f(x)+g(x))dx\\ &= \int_0^1 f(x)dx+\int_0^1 g(x)dx\\ &= T(f)+T(g).\end{aligned}$$

2. Si $\alpha\in\mathbb{R}$ y $f\in C[0,1]$,

$$\begin{aligned} T(\alpha f) &= \int_0^1 (\alpha f)(x)dx\\ &= \int_0^1 \alpha f(x)dx\\ &= \alpha\int_0^1 f(x)dx\\ &= \alpha T(f).\end{aligned}$$ ◀

▶ **Ejemplo 5.8** (Operador identidad) Si $\mathbf{E}$ es un espacio vectorial, claramente $I:\mathbf{E}\to\mathbf{E}$ definido por

$$I(\vec{u})=\vec{u}$$

para todo $\vec{u}\in\mathbf{E}$, es una transformación lineal del espacio $\mathbf{E}$ en sí mismo. A I se le llama el **operador identidad** del espacio $\mathbf{E}$.◀

Propiedades

Sabemos que si A es cualquier conjunto no vacío y $\mathbf{F}$ es cualquier espacio vectorial, entonces el conjunto de funciones con dominio A y valores en $\mathbf{F}$ es un espacio vectorial (cfr. nota 3.7, pág. 138) con las operaciones usuales suma de funciones y multiplicación de un escalar con una función (cfr. definición 3.9, pág. 134). Por ende, el conjunto de todas las transformaciones de un espacio vectorial $\mathbf{E}$ en un espacio vectorial $\mathbf{F}$ con las operaciones usuales de suma de funciones y multiplicación de un escalar con una

función dadas en la definición 3.9 es un espacio vectorial. Ahora supongamos que T_1 y T_2 son un par de transformaciones del espacio $\mathbf{E}$ en el espacio $\mathbf{F}$ que además son lineales y sea $k \in \mathbb{R}$; entonces, si $\vec{x}, \vec{y} \in \mathbf{E}$ y $\alpha \in \mathbb{R}$, se tiene:

1. $$\begin{aligned}(T_1+T_2)(\vec{x}+\vec{y}) &= T_1(\vec{x}+\vec{y})+T_2(\vec{x}+\vec{y})\\ &= T_1(\vec{x})+T_1(\vec{y})+T_2(\vec{x})+T_2(\vec{y})\\ &= T_1(\vec{x})+T_2(\vec{x})+T_1(\vec{y})+T_2(\vec{y})\\ &= (T_1+T_2)(\vec{x})+(T_1+T_2)(\vec{y}).\end{aligned}$$

2. $$\begin{aligned}(T_1+T_2)(\alpha\vec{x}) &= T_1(\alpha\vec{x})+T_2(\alpha\vec{x})\\ &= \alpha T_1(\vec{x})+\alpha T_2(\vec{x})\\ &= \alpha(T_1(\vec{x})+T_2(\vec{x}))\\ &= \alpha(T_1+T_2)(\vec{x}).\end{aligned}$$

3. $$\begin{aligned}(kT_1)(\vec{x}+\vec{y}) &= kT_1(\vec{x}+\vec{y})\\ &= k(T_1(\vec{x})+T_2(\vec{y}))\\ &= kT_1(\vec{x})+kT_2(\vec{y})\\ &= (kT_1)(\vec{x})+(kT_2)(\vec{y}).\end{aligned}$$

4. $$\begin{aligned}(kT_1)(\alpha\vec{x}) &= kT_1(\alpha\vec{x})\\ &= k(\alpha T_1(\vec{x}))\\ &= \alpha(kT_1(\vec{x}))\\ &= \alpha(kT_1)(\vec{x}).\end{aligned}$$

5. Claramente la transformación constante cero, $\theta : \mathbf{E} \to \mathbf{F}$, $\theta(\vec{x}) = \vec{0}_{\mathbf{F}} \quad \forall \vec{x} \in \mathbf{E}$, es lineal.

1 y 2 demuestran que la transformación $T_1 + T_2$ es lineal y 3, 4 prueban que kT_1 es lineal también. Por ende, al adjuntar 5, el subconjunto de transformaciones lineales del espacio $\mathbf{E}$ en el espacio $\mathbf{F}$ es un subespacio vectorial del espacio de funciones de $\mathbf{E}$ en $\mathbf{F}$. Con esto hemos probado el siguiente teorema.

Teorema 5.1 ***(Espacio de transformaciones lineales)*** *Sean* $\mathbf{E}$ *y* $\mathbf{F}$ *un par de espacios vectoriales y sea*

$$\mathscr{L}(\mathbf{E},\mathbf{F}) = \{\, T : \mathbf{E} \to \mathbf{F} \mid T \textit{ es una transformación lineal} \,\},$$

entonces $\mathscr{L}(\mathbf{E},\mathbf{F})$ *es un espacio vectorial con las operaciones usuales de suma de funciones y multiplicación de un escalar con una función.*

Teorema 5.2 ***(Propiedades básicas de transformaciones lineales)*** *Sean* $\mathbf{E}$ *y* $\mathbf{F}$ *un par de espacios vectoriales y* $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$*, entonces:*

1. $T(\vec{0}_{\mathbf{E}}) = \vec{0}_{\mathbf{F}}$.
2. $T(\alpha\vec{x}+\beta\vec{y}) = \alpha T(\vec{x}) + \beta T(\vec{y}) \qquad \forall \vec{x}, \vec{y} \in \mathbf{E}, \forall \alpha, \beta \in \mathbb{R}$.
3. $T(-\vec{x}) = -T(\vec{x}) \qquad \forall \vec{x} \in \mathbf{E}$.
4. $T(\vec{x}-\vec{y}) = T(\vec{x}) - T(\vec{y}) \qquad \forall \vec{x}, \vec{y} \in \mathbf{E}$.

DEMOSTRACIÓN ■ 1.

$$\begin{aligned} T(\vec{0}_{\mathbf{E}}) &= T(0\cdot\vec{0}_{\mathbf{E}}) \\ &= 0\cdot T(\vec{0}_{\mathbf{E}}) \\ &= \vec{0}_{\mathbf{F}}. \end{aligned}$$

2.

$$\begin{aligned} T(\alpha\vec{x}+\beta\vec{y}) &= T(\alpha\vec{x})+T(\beta\vec{y}) \\ &= \alpha T(\vec{x})+\beta T(\vec{y}). \end{aligned}$$

3.

$$\begin{aligned} T(-\vec{x}) &= T((-1)\vec{x}) \\ &= (-1)T(\vec{x}) \\ &= -T(\vec{x}). \end{aligned}$$

4.

$$\begin{aligned} T(\vec{x}-\vec{y}) &= T(\vec{x}+(-\vec{y})) \\ &= T(\vec{x})+T(-\vec{y}) \\ &= T(\vec{x})-T(\vec{y}). \quad ■ \end{aligned}$$

Transformaciones matriciales y representación matricial de una transformación $T \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$

Vimos en los ejemplos 5.2 y 5.3 que las transformaciones lineales de $\mathbb{R}$ en $\mathbb{R}$ son las funciones de la forma $T(x) = ax$, donde a es una constante. El símil para transformaciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ está contenido en el siguiente ejemplo.

► **Ejemplo 5.9** (Transformaciones matriciales) Si $A \in \mathfrak{M}_{m\times n}$, sea $T_A : \mathbb{R}^n \to \mathbb{R}^m$ definida para cada $\vec{x} \in \mathbb{R}^n$ por $T_A(\vec{x}) = A\vec{x}$. Entonces T es lineal; pues $\forall \vec{x}, \vec{y} \in \mathbb{R}^n$ y $\forall \alpha \in \mathbb{R}$:

$$\begin{aligned} T_A(\vec{x}+\vec{y}) &= A(\vec{x}+\vec{y}) \\ &= A\vec{x}+A\vec{y} \\ &= T_A(\vec{x})+T_A(\vec{y}) \quad \text{y} \\ T_A(\alpha\vec{x}) &= A(\alpha\vec{x}) \\ &= \alpha(A\vec{x}) \\ &= \alpha T_A(\vec{x}). ◄ \end{aligned}$$

► **Ejemplo 5.10** Dar una transformación lineal de $\mathbb{R}^3$ en $\mathbb{R}^4$.◄

Solución Sea

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 0 & -1 & 3 \\ 3 & -1 & 4 \\ 2 & 1 & 2 \end{bmatrix} \in \mathfrak{M}_{4\times 3}$$

y $T_A : \mathbb{R}^3 \to \mathbb{R}^4$ definida por

$$T_A(x,y,z) = \begin{bmatrix} 1 & -1 & 2 \\ 0 & -1 & 3 \\ 3 & -1 & 4 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = (x-y+2z, -y+3z, 3x-y+4z, 2x+y+2z).$$

Por el ejemplo anterior T_A es lineal. ✔

Sabemos, del ejemplo 5.9, que toda transformación matricial T_A es lineal. Recíprocamente, si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal, ¿existe una matriz $A \in \mathfrak{M}_{m\times n}$ tal que $T = T_A$? La respuesta es por lo menos afirmativa cuando $m = n = 1$ para el caso de transformaciones lineales de $\mathbb{R}$ en $\mathbb{R}$ como probamos en el ejemplo 5.3 (pág. 417). Esto nos induce a intentar probar el caso general. Sean

$$\vec{e}_i = (\underbrace{0,\ldots,0,\overset{i}{1},0,\ldots,0}_{n}), \quad i = 1,2,\ldots,n,$$

la base canónica de $\mathbb{R}^n$ y

$$\vec{f}_j = (\underbrace{0,\ldots,0,\overset{j}{1},0,\ldots,0}_{m}), \quad j = 1,2,\ldots,m,$$

la base canónica de $\mathbb{R}^m$. Entonces, si $\vec{u} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$,

$$\vec{u} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n;$$

así que

$$\begin{aligned} T(\vec{u}) &= T(x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n) \\ &= T(x_1\vec{e}_1) + T(x_2\vec{e}_2) + \cdots + T(x_n\vec{e}_n) \\ &= x_1T(\vec{e}_1) + x_2T(\vec{e}_2) + \cdots + x_nT(\vec{e}_n). \end{aligned}$$

Sea A la matriz $m \times n$ que tiene por columnas a los vectores $T(\vec{e}_i)$, $i = 1,2,\ldots,n$, descritos de acuerdo con la base canónica $\{\vec{f}_j\}$. Entonces

$$\begin{aligned} T(\vec{u}) &= x_1T(\vec{e}_1) + x_2T(\vec{e}_2) + \cdots + x_nT(\vec{e}_n) \\ &= \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \cdots & T(\vec{e}_n) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A\vec{u} \end{aligned}$$

(cfr. ejemplo 1.15, pág. 12); luego $T = T_A$. Hemos probado así el siguiente teorema.

Teorema 5.3 *Sean $\{\vec{e}_i\}$, $\{\vec{f}_j\}$ las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente, y $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Entonces existe una matriz $A \in \mathfrak{M}_{m\times n}$ tal que $T = T_A$. De hecho, una matriz A que sirve para este propósito es aquella cuyas columnas son los vectores $T(\vec{e}_i)$, $i = 1,\ldots,n$, descritos en la base canónica de $\mathbb{R}^m$.*

❍ **Nota 5.1**

1. El teorema 5.3 significa que las únicas transformaciones lineales que existen entre espacios $\mathbb{R}^k$ son las transformaciones matriciales T_A. Éste es un resultado muy sencillo, pero también muy importante, que el lector debe tener siempre presente.
2. La matriz A del teorema 5.3 no es única en el sentido de que se pueden tener resultados análogos a este teorema utilizando otras bases (cfr. sección 5.2).

Definición 5.2 *Sean $T \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$, $\{\vec{e}_i\}$ y $\{\vec{f}_j\}$ las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente. A la matriz $A \in \mathfrak{M}_{m\times n}$ cuyas columnas son los vectores $T(\vec{e}_i)$, $i = 1, 2, \ldots, n$, descritos en la base canónica $\{\vec{f}_j\}$, se le dice la* ***representación matricial*** *de la transformación lineal[2] T relativa a las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$.*

El teorema 5.3 también es muy útil para determinar si una transformación $T : \mathbb{R}^n \to \mathbb{R}^m$ es lineal o no.[3] Para ello basta tomar la matriz A como antes y comprobar si $T(\vec{u}) = A\vec{u}\ \forall \vec{u} \in \mathbb{R}^n$; pues en caso afirmativo $T = T_A$ y, por ende, T será lineal; en caso contrario, por el teorema precedente, T no puede ser lineal.

► **Ejemplo 5.11** Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ la transformación definida por $T(x, y) = (-x+2y, x, 2x-y)$. Mostrar que T es lineal y encontrar la representación matricial de T relativa a las bases canónicas. ◄

DEMOSTRACIÓN ■

$$\begin{aligned} T(1,0) &= (-1,1,2),\\ T(0,1) &= (2,0,-1). \end{aligned}$$

Si

$$A = \begin{bmatrix} -1 & 2 \\ 1 & 0 \\ 2 & -1 \end{bmatrix}$$

se tiene, para todo $\vec{u} = (x, y) \in \mathbb{R}^2$,

$$\begin{aligned} A\vec{u} &= \begin{bmatrix} -1 & 2 \\ 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} -x+2y \\ x \\ 2x-y \end{bmatrix} = T(x,y). \end{aligned}$$

Esto es

$$T = T_A.$$

Lo cual prueba que T es lineal y A es entonces la representación matricial de T (relativa a las bases canónicas). ■

5.1.2 Núcleo e imagen de una transformación lineal

Funciones inversas

Sean A y B un par de conjuntos no vacíos y $f : A \to B$. Una función $g : B \to A$ es **función inversa** de f si

$$\begin{aligned} f(g(y)) &= y\ \ \forall y \in B \text{ y}\\ g(f(x)) &= x\ \ \forall x \in A. \end{aligned}$$

[2] Más adelante, en la sección 5.2, extenderemos este concepto a transformaciones $T \in \mathscr{L}(\mathbf{E}, \mathbf{F})$ en espacios de dimensión finita.

[3] Lo cual es generalmente laborioso y tedioso aunque no difícil de hacer.

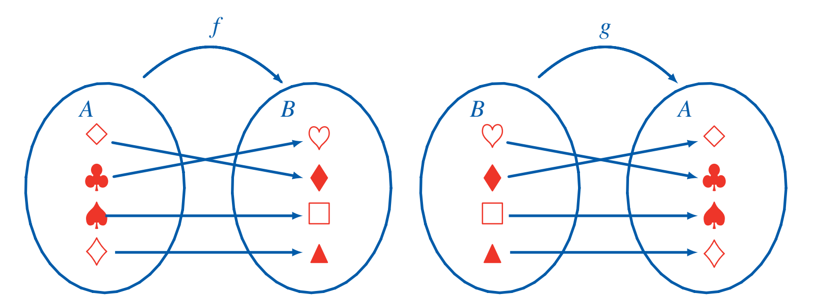


Figura 5-1 • Diagrama de flechas que define un par de funciones. La función f queda definida mediante la asignación que a cada elemento del conjunto A se hace con un elemento del conjunto B por medio de la flecha que se indica; $f(\diamondsuit) = \blacklozenge$, $f(\clubsuit) = \heartsuit$, etc. De manera análoga queda definida la función g.

En tal caso, se dice que la función f es **invertible**. Por ejemplo, si $A = [0,2]$, $B = [0,4]$ y $f : A \to B$ se define por $f(x) = x^2$; sea $g : B \to A$ definida por $g(y) = \sqrt{y}$, entonces g es función inversa de f. En efecto:

$$f(g(y)) = f(\sqrt{y}) = y$$

y

$$g(f(x)) = g(x^2) = x$$

para todo $y \in B = [0,4]$ y para todo $x \in A = [0,2]$. De existir una inversa de una función ésta es única, la demostración de este hecho se deja como ejercicio al lector. Entonces, si f tiene inversa ésta se denota por f^{-1} y se le llama la **función inversa** de f (por ser única, la notación y el artículo están justificados). En la figura 5-1 se bosquejan dos diagramas de flechas que definen sendas funciones $f : A \to B$ y $g : B \to A$. Claramente, se tiene $g = f^{-1}$. De la misma figura se observa que, en este caso, para cada $y \in B$ existe un único $x \in A$ tal que $f(x) = y$, entonces la función $f^{-1}(y) = x$ está bien definida y coincide con la función g; luego f es una función invertible. Así, para que una función sea invertible, es necesario y suficiente que para cada elemento del contradominio exista un único elemento del dominio cuya imagen, bajo la función, sea dicho elemento. Estas características están especificadas en las siguientes definiciones.

Definición 5.3 ***(Funciones inyectivas)*** *Sean A y B un par de conjuntos y $f : A \to B$. La función f es* ***inyectiva*** *(uno a uno) si elementos distintos del dominio tienen imágenes distintas; esto es,*

$$x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$$

para $x_1, x_2 \in A$; o equivalentemente,

$$f(x_1) = f(x_2) \Rightarrow x_1 = x_2.$$

► **Ejemplo 5.12** Sea $f : \mathbb{R} \to \mathbb{R}$ la función definida por $f(x) = x^3$, para cada $x \in \mathbb{R}$. Si $x_1, x_2 \in \mathbb{R}$ y $x_1 \neq x_2$, entonces

$$x_1^2 + x_1x_2 + x_2^2 \neq 0. \quad (5.1)$$

En efecto, si uno de x_1, x_2 es cero, el otro debe ser distinto de cero (pues $x_1 \neq x_2$); y por tanto, $x_1^3 \neq x_2^3$; así que podemos suponer que x_1 y x_2 son ambos distintos de cero; por tanto, ya que la ecuación cuadrática (5.1) en x_1 tiene discriminante $-3x_2^2 < 0$, se tiene que ésta no tiene raíces reales; por lo que

$$x_1^3 - x_2^3 = (x_1^2 + x_1x_2 + x_2^2)(x_1 - x_2) \neq 0.$$

Luego,

$$x_1 \neq x_2 \Rightarrow x_1^3 \neq x_2^3;$$

por tanto, la función $f(x) = x^3$ es inyectiva. De manera gráfica también se puede observar que la función f es inyectiva.◀

Definición 5.4 *Se dice que una función $f : A \to B$ es **suprayectiva** (sobre) si para todo $y \in B$ existe $x \in A$ tal que $f(x) = y$. Si f es una función suprayectiva, se acostumbra decir que f* es una función de A sobre B *en lugar de f* es una función de A en B.

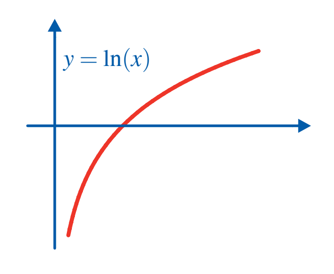


Figura 5-2 •

▶ **Ejemplo 5.13** De la gráfica de la función $f(x) = \ln x$, figura 5-2, se puede ver que ésta es una función suprayectiva de $\mathbb{R}$ sobre $\mathbb{R}$. De hecho, si $y \in \mathbb{R}$, entonces $x = e^y$ satisface $\ln(x) = y$, como el lector puede comprobar utilizando las propiedades del logaritmo natural y la exponencial.◀

Definición 5.5 *$f : A \to B$ es una función **biyectiva** si f es inyectiva y suprayectiva.*

▶ **Ejemplo 5.14** La función $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = \ln(x)$, es inyectiva y suprayectiva (cfr. figura 5-2), luego f es biyectiva.◀

Resumimos, en el siguiente teorema, los resultados precedentes.

Teorema 5.4 *Sea $f : A \to B$. Entonces f es invertible si y sólo si f es biyectiva.*

❍ **Nota 5.2** Observe que $f : A \to B$ es biyectiva si y sólo si $f^{-1} : B \to A$ es biyectiva, y que en tal caso $(f^{-1})^{-1} = f$.

▶ **Ejemplo 5.15** La función $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = \ln(x)$, es biyectiva y $f^{-1}(x) = e^x$.◀

Núcleo

En el caso de una transformación lineal de un espacio vectorial en otro, la inyectividad es una característica muy fácil de verificar mediante lo que llamaremos el núcleo de dicha aplicación; mientras que la suprayectividad será también una propiedad sencilla de comprobar mediante la imagen de la transformación. Desarrollamos estos conceptos a continuación.

Definición 5.6 *Si $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$, se define el* ***núcleo*** *(**kernel**) de T como:*

$$\operatorname{Ker}(T) = \left\{ \vec{x} \in \mathbf{E} \,|\, T(\vec{x}) = \vec{0}_{\mathbf{F}} \right\}.$$

► **Ejemplo 5.16** Si $T \in \mathscr{L}(\mathbb{R}^3,\mathbb{R}^2)$ está definida por[4]

$$T(x_1,x_2,x_3) = (x_1 - 2x_2, x_2 + 3x_3).$$

Entonces $\vec{x} = (x_1,x_2,x_3) \in \operatorname{Ker}(T)$ si y sólo si $T(x_1,x_2,x_3) = (0,0)$. Luego se debe tener

$$\begin{aligned} x_1 - 2x_2 &= 0 \\ x_2 + 3x_3 &= 0. \end{aligned}$$

Resolvamos el sistema homogéneo escalonado

$$\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 3 \end{bmatrix}$$

haciendo sustitución regresiva:
$$\begin{aligned} x_2 &= -3x_3 \\ x_1 &= 2x_2 = -6x_3. \end{aligned}$$

De donde

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -6r \\ -3r \\ r \end{bmatrix} = r \begin{bmatrix} -6 \\ -3 \\ 1 \end{bmatrix}.$$

Así,

$$\operatorname{Ker}(T) = \{ \vec{u} \in \mathbb{R}^3 \,|\, \vec{u} = (-6r,-3r,r),\ r \in \mathbb{R} \}.$$ ◄

► **Ejemplo 5.17** (Núcleo de una transformación matricial) Sean $A \in \mathfrak{M}_{m\times n}$ y la transformación matricial $T_A : \mathbb{R}^n \to \mathbb{R}^m$, $T_A(\vec{x}) = A\vec{x}$. Entonces $\vec{x} \in \operatorname{Ker}(T) \Leftrightarrow A\vec{x} = \vec{0}_{\mathbb{R}^m} \Leftrightarrow \vec{x}$ está en el espacio nulo de la matriz A; esto es, en el espacio solución del sistema homogéneo $A\vec{x} = \vec{0}$. Es decir, el núcleo de una transformación matricial T_A es el espacio nulo[5] de la matriz A. ◄

► **Ejemplo 5.18** Sea $\mathbf{E} = C^1[0,1]$ y $T : \mathbf{E} \to C[0,1]$, la transformación lineal $T(f) = f'$ (cfr. ejemplo 5.4). Entonces

$$\begin{aligned} \operatorname{Ker}(T) &= \{ f \in C^1[0,1] \,|\, T(f) = 0 \} \\ &= \{ f \in C^1[0,1] \,|\, f' = 0 \} \\ &= \{ f \in C^1[0,1] \,|\, f'(x) = 0\ \forall x \in [0,1] \} \\ &= \{ f \in C^1[0,1] \,|\, f \text{ es constante en } [0,1] \}. \end{aligned}$$ ◄

Sean ahora $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$, $\vec{x}_1,\vec{x}_2 \in \operatorname{Ker}(T)$, y $\alpha,\beta \in \mathbb{R}$; entonces $T(\vec{x}_1) = T(\vec{x}_2) = \vec{0}_{\mathbf{F}}$. Luego

$$T(\alpha\vec{x}_1 + \beta\vec{x}_2) = \alpha T(\vec{x}_1) + \beta T(\vec{x}_2) = \vec{0}_{\mathbf{F}} + \vec{0}_{\mathbf{F}} = \vec{0}_{\mathbf{F}},$$

[4] Queda como ejercicio para el lector comprobar que efectivamente T es una transformación lineal (cfr. ejemplo 5.5 y la definición de transformación lineal o el ejemplo 5.11 y la discusión que lo precede).

[5] Cfr. ejemplo 3.20, página 141.

por lo que $\alpha\vec{x}_1+\beta\vec{x}_2\in \text{Ker}(T)$; y puesto que[6] $T(\vec{0}_{\mathbf{E}})=\vec{0}_{\mathbf{F}}$, se sigue que el núcleo de toda transformación lineal es un subespacio vectorial y por ende un espacio vectorial. Resumimos este resultado en el siguiente teorema.

Teorema 5.5 *Si* $\mathscr{L}(\mathbf{E},\mathbf{F})$, *entonces*[7] $\text{Ker}(T)<\mathbf{E}$.

En el teorema 5.6 se ve la relación entre la inyectividad y el núcleo de una transformación lineal que acabamos de mencionar.

Teorema 5.6 ***(Núcleo e inyectividad)*** *Si* $\mathbf{E}$ *y* $\mathbf{F}$ *son espacios vectoriales y* $T\in\mathscr{L}(\mathbf{F},\mathbf{F})$, *entonces* T *es inyectiva si y sólo si* $\text{Ker}(T)=\{\vec{0}_{\mathbf{E}}\}$.

DEMOSTRACIÓN ■

1. Supongamos que T es inyectiva. Si $\vec{x}\in\text{Ker}(T)$, $T(\vec{x})=\vec{0}_{\mathbf{F}}$ y puesto que $T(\vec{0}_{\mathbf{E}})=\vec{0}_{\mathbf{F}}$, $\vec{x}=\vec{0}_{\mathbf{E}}$; por tanto, $\text{Ker}(T)=\{\vec{0}_{\mathbf{E}}\}$.
2. Supongamos ahora que $\text{Ker}(T)=\{\vec{0}_{\mathbf{E}}\}$. Si $\vec{x},\vec{y}\in\mathbf{E}$ y $T(\vec{x})=T(\vec{y})$, entonces $T(\vec{x}-\vec{y})=T(\vec{x})-T(\vec{y})=\vec{0}_{\mathbf{F}}$. Por tanto, $\vec{x}-\vec{y}\in\text{Ker}(T)=\{\vec{0}_{\mathbf{E}}\}$. Así que $\vec{x}-\vec{y}=\vec{0}_{\mathbf{E}}$; es decir, $\vec{x}=\vec{y}$; lo cual prueba que T es inyectiva. ■

► **Ejemplo 5.19** Si $A\in\mathfrak{M}_{n\times n}$, la transformación matricial T_A de $\mathbb{R}^n$ en $\mathbb{R}^n$ es inyectiva si y sólo si A es invertible. En efecto:

$$\begin{aligned} T_A \text{es inyectiva} &\Leftrightarrow \text{Ker}(T_A)=\{\vec{0}_{\mathbb{R}^n}\}\\ &\Leftrightarrow \text{la única solución de } A\vec{x}=\vec{0}_{\mathbb{R}^n} \text{ es } \vec{x}=\vec{0}_{\mathbb{R}^n}\\ &\Leftrightarrow A \text{ es invertible.} \end{aligned}$$ ◄

Definición 5.7 *Si* $T\in\mathscr{L}(\mathbf{E},\mathbf{E})$*; es decir,* T *es una transformación lineal de un espacio* $\mathbf{E}$ *en sí mismo, se acostumbra decir que* T *es un* ***operador lineal*** *en el espacio* $\mathbf{E}$.

Más adelante veremos que a todo operador lineal T en un espacio $\mathbf{E}$ de dimensión finita le corresponde una transformación matricial T_A; de tal suerte que el operador se puede evaluar por medio de esta transformación matricial. Por esta razón y el ejemplo precedente, es común decir que un operador lineal es **no singular** cuando es invertible. En este texto utilizaremos indistintamente los calificativos invertible y no singular cuando un operador lineal sea invertible.

► **Ejemplo 5.20** Sea $A=\begin{bmatrix}1 & -1\\ 2 & 1\end{bmatrix}$ y $T_A:\mathbb{R}^2\to\mathbb{R}^2$ la transformación matricial definida por $T_A(x,y)=A\begin{bmatrix}x\\ y\end{bmatrix}$; puesto que

$$A=\begin{bmatrix}1 & -1\\ 2 & 1\end{bmatrix}\sim\begin{bmatrix}1 & -1\\ 0 & 3\end{bmatrix}\sim I_2,$$

se deduce que A es invertible, luego el operador lineal T_A es inyectivo. ◄

[6] Cfr. teorema 5.2, página 419.

[7] Recuerde que la notación $S<\mathbf{E}$ significa que S es un subespacio vectorial de $\mathbf{E}$ (cfr. definición 3.10, pág. 139).

▶ **Ejemplo 5.21** Sea $T:\mathbb{R}^2\to\mathbb{R}^3$ definida por $T(x,y)=(2x,3x+4,x+y)$, se deja como ejercicio al lector comprobar que T es lineal. Determinar si T es inyectiva.◀

Solución

$$\begin{aligned} T(x,y) &= (0,0,0) \Rightarrow \\ (2x,3x+y,x+y) &= (0,0,0) \Rightarrow \\ x &= 0, \quad y=0. \end{aligned}$$

Por tanto, $\text{Ker}(T)=\{(0,0)\}$, por lo que T es inyectiva. ✔

▶ **Ejemplo 5.22** Sea $T:\mathbb{R}^3\to\mathbb{R}^3$ definida por $T(x,y,z)=(z,0,x+y)$. Sean $(x_1,y_1,z_1),(x_1,y_2,z_3)\in\mathbb{R}^3$ y $\alpha\in\mathbb{R}$, entonces

$$\begin{aligned} T\left((x_1,y_1,z_1)+(x_2,y_2,z_2)\right) &= T(x_1+x_2,y_1+y_2,z_1+z_2) \\ &= (z_1+z_2,0,x_1+x_2+y_1+y_2) \\ &= (z_1,0,x_1+y_1)+(z_2,0,x_2+y_2) \\ &= T(x_1,y_1,z_1)+T(x_2,y_2,z_2) \end{aligned}$$

y

$$\begin{aligned} T(\alpha(x_1,y_1,z_1)) &= T(\alpha x_1,\alpha y_1,\alpha z_1) \\ &= (\alpha z_1,0,\alpha x_1+\alpha y_1) \\ &= \alpha(z_1,0,x_1+y_1) \\ &= \alpha T(x_1,y_1,z_1). \end{aligned}$$

Lo cual demuestra que T es un operador lineal. Por otra parte, si $\vec{u}=(-1,1,0)$, entonces

$$T(\vec{u})=(0,0,-1+1)=(0,0,0),$$

por lo que $\vec{u}\in\text{Ker}(T)$ y por tanto $\text{Ker}(T)\neq\{(0,0,0)\}$. Luego el operador lineal T no es inyectivo.◀

Imagen de una transformación

Definición 5.8 *Sean* $\mathbf{E}$, $\mathbf{F}$ *un par de espacios vectoriales,* $T\in\mathscr{L}(\mathbf{E},\mathbf{F})$ *y* $S\subset\mathbf{E}$ *un subconjunto no vacío:*

1. *Al conjunto*

$$T(S)=\{\,T(\vec{x})\,|\,\vec{x}\in S\,\}=\{\,\vec{y}\in\mathbf{F}\,|\,\text{existe } \vec{x}\in S \text{ con } T(\vec{x})=\vec{y}\,\}$$

se le llama la ***imagen*** *de* S *bajo* T*. También se dice que* S *se transforma en* $T(S)$ *bajo la aplicación* T*.*

2. *En particular,* $T(\mathbf{E})$ *se llama la* **imagen** *de la transformación* T*.*

▶ **Ejemplo 5.23** Sea $T:\mathbb{R}^2\to\mathbb{R}^2$ definida por $T(x,y)=(2x,y)$ para todo $(x,y)\in\mathbb{R}^2$. Puesto que

$$T(x,y)=\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix},$$

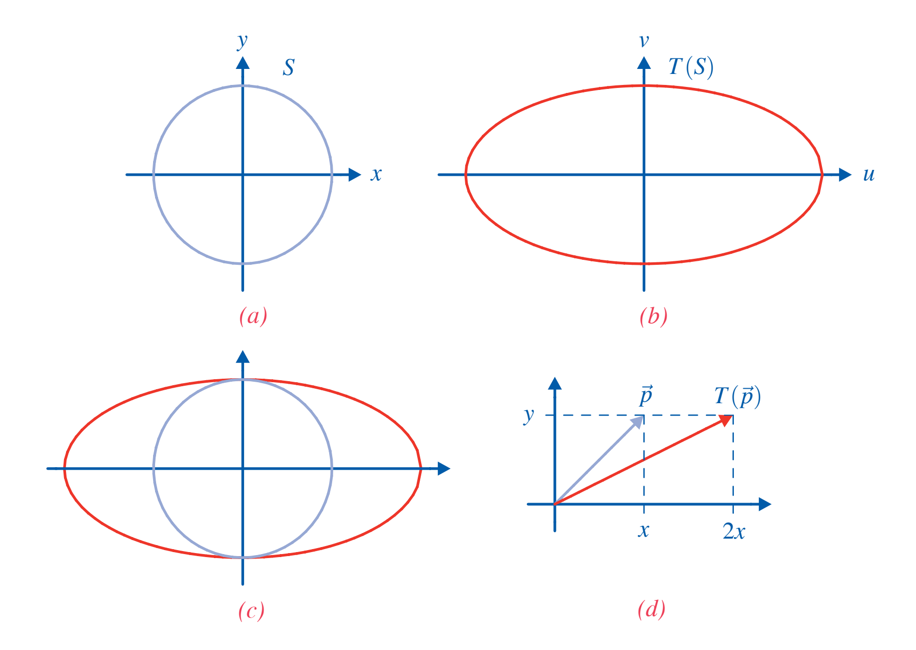


Figura 5-3 •

es inmediato que T es una transformación lineal. Sea

$$S = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\};$$

i.e., S es la circunferencia con centro en el origen y radio 1. El objetivo es determinar geométricamente en qué se transforma el conjunto S bajo T; es decir, $T(S)$. Para ello podemos bosquejar el plano $\mathbb{R}^2$ y una copia del mismo donde, para mejor comprensión, hemos elegido la letra u en el eje de las abscisas y la letra v para el eje de las ordenadas en lugar de x y y, respectivamente, del plano original, para graficar la imagen de S bajo la transformación T, como se ilustra en la figura 5-3*(a)* y *(b)*. Entonces, si

$$(u,v) = T(x,y) = (2x,y)$$

se tiene

$$u = 2x,$$
$$v = y.$$

Luego, si $(x,y) \in S$,

$$\left(\frac{u}{2}\right)^2 + v^2 = x^2 + y^2 = 1.$$

De esta manera todo punto (x,y) de la circunferencia S (cfr. figura 5.23*(a)*) se transforma en un punto $T(x,y) = (u,v)$ de la elipse $(u^2/4) + v^2 = 1$ bosquejada en la figura 5-3*(b)*. Por tanto, la imagen de la circunferencia S, bajo la transformación T, es la elipse $T(S)$ con ecuación, en el plano u, v, $(u^2/4) + v^2 = 1$. Naturalmente, es posible utilizar el mismo plano para bosquejar S y su imagen $T(S)$ (figura 5-3*(c)*). En general, esta aplicación transforma un vector $\vec{p}$ en otro vector $T(\vec{p})$ que sufre una dilatación, en un factor de 2, en la abscisa y que permanece invariante en la ordenada (figura 5-3*(d)*).◀

Teorema 5.7 *Sean* $\mathbf{E}$, $\mathbf{F}$ *espacios vectoriales y* $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$*:*

1. *Si* $S < \mathbf{E}$*, entonces* $T(S) < \mathbf{F}$.
2. $T(\mathbf{E}) < \mathbf{F}$.
3. *Si* $S = \text{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$*, entonces* $T(S) = \text{gn}(T(\vec{u}_1), T(\vec{u}_2), \ldots, T(\vec{u}_n))$.

DEMOSTRACIÓN ■ 1. *(a)* Como $S < \mathbf{E}$, $\vec{0}_{\mathbf{E}} \in S$ y entonces, al ser T lineal, $T(\vec{0}_{\mathbf{E}}) = \vec{0}_{\mathbf{F}}$; luego $\vec{0}_{\mathbf{F}} \in T(S)$.

(b) Si $\vec{y}_1, \vec{y}_2 \in T(S)$, existen $\vec{x}_1, \vec{x}_2 \in S$ tales que $T(\vec{x}_1) = \vec{y}_1$ y $T(\vec{x}_2) = \vec{y}_2$; entonces, dado que $\vec{x}_1 + \vec{x}_2 \in S$ (porque $S < \mathbf{E}$), $\vec{y}_1 + \vec{y}_2 = T(\vec{x}_1) + T(\vec{x}_2) = T(\vec{x}_1 + \vec{x}_2) \in T(S)$.

(c) Sea $\vec{y} \in T(S)$ y $\alpha \in \mathbb{R}$, entonces existe $\vec{x} \in \mathbf{E}$ tal que $\vec{y} = T(\vec{x})$; dado que $\alpha\vec{x} \in S$ (porque $S < \mathbf{E}$), $\alpha\vec{y} = \alpha T(\vec{x}) = T(\alpha\vec{x}) \in T(S)$.

De *(a)*, *(b)* y *(c)* se tiene que $T(S) < \mathbf{F}$.

2. Es consecuencia inmediata del inciso 1.

3. Si $\vec{v} \in \text{gn}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$, existen $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ tales que

$$\vec{v} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_n\vec{u}_n,$$

entonces

$$\begin{aligned} T(\vec{v}) &= T(\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_n\vec{u}_n) \\ &= T(\alpha_1\vec{u}_1) + T(\alpha_2\vec{u}_2) + \cdots + T(\alpha_n\vec{u}_n) \\ &= \alpha_1 T(\vec{u}_1) + \alpha_2 T(\vec{u}_2) + \cdots + \alpha_n T(\vec{u}_n); \end{aligned}$$

de donde $T(S) = \text{gn}(T(\vec{u}_1), T(\vec{u}_2), \ldots, T(\vec{u}_n))$. ■

▶ **Ejemplo 5.24** Sea $T; \mathbb{R}^3 \to \mathbb{R}^3$ la transformación definida como

$$T(x,y,z) = (x + 2y - z, y + z, x + y - 2z).$$

1. Encontrar una matriz $A \in \mathfrak{M}_{3\times 3}$ tal que $T(\vec{u}) = A\vec{u}\ \ \forall \vec{u} \in \mathbb{R}^3$.
2. Demostrar que T es un operador lineal en $\mathbb{R}^3$.
3. Hallar una base y $\dim(T(\mathbb{R}^3))$.
4. Hallar una base de $\text{Ker}(T)$ y $\dim(\text{Ker}(T))$.
5. ¿Cuál es el valor de $\dim(\text{Ker}(T)) + \dim((T(\mathbf{E}))$?◀

Solución 1.

$$\begin{aligned} T(\vec{e}_1) &= T(1,0,0) = (1,0,1), \\ T(\vec{e}_2) &= T(1,0,0) = (2,1,1), \\ T(\vec{e}_3) &= T(0,0,1) = (-1,1,-2). \end{aligned}$$

Sea $A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$, entonces $\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + 2y - z \\ y + z \\ x + y - 2z \end{bmatrix} = T(x,y,z)$.

Esto es, $T(\vec{u}) = A\vec{u} \quad \forall \vec{u} \in \mathbb{R}^3$.

2. Puesto que $T = T_A$, se sigue que T es lineal.

3. Sabemos que $\vec{e}_1 = (1,0,0), \vec{e}_2 = (0,1,0), \vec{e}_3 = (0,0,1)$ forman una base de $\mathbb{R}^3$; así $T(\vec{e}_1), T(\vec{e}_2), T(\vec{e}_3)$ generan a $T(\mathbb{R}^3)$.

   Entonces,

$$\begin{aligned} T(\mathbb{R}^3) &= \text{gn}(T(\vec{e}_1), T(\vec{e}_2), T(\vec{e}_3)) \\ &= \text{gn}((1,0,1),(2,1,1),(-1,1,-2)). \end{aligned}$$

   Ahora extraigamos una base para $T(\mathbb{R}^3)$ a partir de estos generadores (cfr. teorema 3.18 y ejemplo 3.61, pág. 166) llevando la matriz A a forma escalonada.

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

   Una base para $T(\mathbb{R}^3)$ es entonces $\mathscr{B} = \{(1,0,1),(2,1,1)\}$ y $\dim(T(\mathbb{R}^3)) = 2$.

4. Para hallar el núcleo de T tenemos

$$\begin{aligned} T(x,y,z) = (0,0,0) &\Leftrightarrow \\ (x-2y-z, y+z, x+y-2z) &= (0,0,0); \end{aligned}$$

   que equivale a resolver el sistema homogéneo

$$\begin{array}{rcrcrcl} x & - & 2y & - & z & = & 0 \\ & & y & + & z & = & 0 \\ x & + & y & - & 2z & = & 0 \end{array}$$

   Resolvamos este sistema llevando la matriz de coeficientes a forma escalonada:

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

   y después haciendo sustitución regresiva:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3r \\ -r \\ r \end{bmatrix} = r \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}.$$

   Por tanto,

$$\text{Ker}(T) = \text{gn}((3,-1,1)).$$

   Entonces, una base para $\text{Ker}(T)$ es $\mathscr{B} = \{(3,-1,1)\}$ y $\dim(\text{Ker}(T)) = 1$.

5. $$\dim(\text{Ker}(T)) + \dim((T(\mathbf{E})) = 1 + 2 = 3 = \dim(\mathbb{R}^3). \quad ✔$$

El precedente ejemplo motiva los resultados contenidos en los teoremas 5.8 y 5.9.

Teorema 5.8 *Si* $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$ *y* $\dim(\mathbf{E}) = n$*, entonces*

$$\dim(T(\mathbf{E})) + \dim(\mathrm{Ker}(T)) = n \tag{5.2}$$

DEMOSTRACIÓN ■ Sea $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k\}$ una base de $\mathrm{Ker}(T)$. Completemos ésta a una base

$$\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k, \vec{f}_1, \vec{f}_2, \ldots, \vec{f}_{n-k}\}$$

de $\mathbf{E}$ (cfr. el procedimiento para este fin dado en la pág. 167). Entonces, si $\vec{u} = a_1\vec{e}_1 + \cdots + a_k\vec{e}_k + b_1\vec{f}_1 + \cdots + b_{n-k}\vec{f}_{n-k} \in \mathbf{E}$,

$$\begin{aligned} T(\vec{u}) &= a_1T(\vec{e}_1) + \cdots + a_kT(\vec{e}_k) + b_1T(\vec{f}_1) + \cdots + b_{n-k}T(\vec{f}_{n-k}) \\ &= \vec{0}_{\mathbf{F}} + \cdots + \vec{0}_{\mathbf{F}} + b_1T(\vec{f}_1) + \cdots + b_{n-k}T(\vec{f}_{n-k}) \\ &= b_1T(\vec{f}_1) + \cdots + b_{n-k}T(\vec{f}_{n-k}); \end{aligned}$$

de donde $T(\mathbf{E}) = \mathrm{gn}(T(\vec{f}_1), \ldots, T(\vec{f}_{n-k}))$. Sean $\beta_1, \ldots, \beta_{n-k} \in \mathbb{R}$ tales que

$$\beta_1T(\vec{f}_1) + \cdots + \beta_{n-k}T(\vec{f}_{n-k}) = \vec{0}_{\mathbf{F}},$$

entonces

$$\begin{aligned} T(\beta_1\vec{f}_1 + \cdots + \beta_{n-k}\vec{f}_{n-k}) &= \beta_1T(\vec{f}_1) + \cdots + \beta_{n-k}T(\vec{f}_{n-k}) \\ &= \vec{0}_{\mathbf{F}}. \end{aligned}$$

Por lo que $\beta_1\vec{f}_1 + \cdots + \beta_{n-k}\vec{f}_{n-k} \in \mathrm{Ker}(T)$; por tanto existen $-\alpha_1, \ldots, -\alpha_k \in \mathbb{R}$ tales que

$$\beta_1\vec{f}_1 + \cdots + \beta_{n-k}\vec{f}_{n-k} = (-\alpha_1)\vec{e}_1 + \cdots + (-\alpha_k)\vec{e}_k$$

(pues $\{\vec{e}_1, \ldots, \vec{e}_k\}$ es una base de $\mathrm{Ker}(T)$). Entonces

$$\alpha_1\vec{e}_1 + \cdots + \alpha_k\vec{e}_k + \beta_1\vec{f}_1 + \cdots + \beta_{n-k}\vec{f}_{n-k} = \vec{0}_{\mathbf{E}}$$

y ya que los vectores $\vec{e}_1, \ldots, \vec{e}_k, \vec{f}_1, \ldots, \vec{f}_{n-k}$ son L.I., se tiene que los escalares $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_{n-k}$ son todos nulos; en particular

$$\beta_1 = \cdots = \beta_{n-k} = 0.$$

Luego los vectores $T(\vec{f}_1), \ldots, T(\vec{f}_{n-k})$ son L.I y, ya que generan a $T(\mathbf{E})$, forman una base de $T(\mathbf{E})$. Así, $\dim(T(\mathbf{E})) = n - k$; por ende,

$$\dim(\mathrm{Ker}(T)) + \dim(T(\mathbf{E})) = n = \dim(\mathbf{E}). \quad ■$$

Definición 5.9 *Si* $\mathbf{E}$ *tiene dimensión finita y* $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$*, a la dimensión del subespacio* $T(\mathbf{E})$ *se le llama el* ***rango*** *de la transformación* T *y se denota por* $\mathrm{Rang}(T)$*; es decir,*

$$\mathrm{Rang}(T) = \dim(T(\mathbf{E})).$$

Teorema 5.9 *Sean $T \in \mathscr{L}(\mathbb{R}^{\mathbf{n}}, \mathbb{R}^{\mathbf{m}})$ y A la representación matricial de T relativa a las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$. Entonces:*

1. $T(\mathbb{R}^n) = E_c(A)$,[8] *esto es, $\vec{v} \in T(\mathbb{R}^n)$ si y sólo si el sistema $A\vec{u} = \vec{v}$ tiene solución.*
2. $\dim(T(\mathbb{R}^n)) = \text{Rang}(A)$, *el rango de la matriz A.*
3. $\vec{u} \in \text{Ker}(T) \Leftrightarrow \vec{u}$ *pertenece al espacio nulo de A; esto es, $\vec{u}$ es solución del sistema homegéneo $A\vec{u} = \vec{0}_{\mathbb{R}^m}$.*
4. $\dim(\text{Ker}(T)) = \text{Nul}(A)$, *la nulidad de A, la dimensión del espacio solución del sistema $A\vec{x} = \vec{0}_{\mathbb{R}^m}$.*
5. *Si $m = n$, el operador lineal T es inyectivo (por tanto biyectivo) si y sólo si A es invertible.*

La demostración de este teorema es sencilla y se deja como ejercicio al lector.

▶ **Ejemplo 5.25** ¿Es posible construir una transformación lineal $T \in \mathscr{L}(\mathbb{R}^4, \mathbb{R}^3)$ que sea inyectiva? Supongamos que sí. Por 5.2 del teorema 5.8 tenemos

$$4 = \dim(T(\mathbb{R}^4)) + \dim(\text{Ker}(T)) = \dim(T(\mathbb{R}^4)) + 0 = \dim(T(\mathbb{R}^4))$$

pues T es inyectiva y por ende $\text{Ker}(T) = \{\vec{0}_{\mathbf{E}}\}$, luego $\dim(\text{Ker}(T)) = 0$. Pero $T(\mathbb{R}^4) < \mathbb{R}^3$ y entonces $\dim(T(\mathbb{R}^4))$ es a lo más 3. Así

$$4 = \dim(T(\mathbb{R}^4)) \leq 3.$$

Lo cual es una contradicción; por tanto es imposible que T sea inyectiva.◀

El mismo razonamiento del ejemplo precedente es válido para cualquier $T \in \mathscr{L}(\mathbf{E}, \mathbf{F})$ donde $\mathbf{E}$ tiene dimensión finita y $\dim(\mathbf{F}) < \dim(\mathbf{E})$ para mostrar que T no puede ser inyectiva.

Como mencionamos más arriba, para el caso de espacios vectoriales y de una transformación lineal entre ellos, hay condiciones sencillas para probar la suprayectividad e inyectividad.

Teorema 5.10 *Sean $\mathbf{E}$ y $\mathbf{F}$ un par de espacios vectoriales y $T : \mathbf{E} \to \mathbf{F}$ una transformación lineal. Entonces:*

1. *T es suprayectiva $\Leftrightarrow T(\mathbf{E}) = \mathbf{F}$.*
2. *Si $\dim(\mathbf{E}) = n < \infty$, T es suprayectiva si y sólo si $\dim(T(\mathbf{E})) = \dim(\mathbf{F})$.*
3. *Si $\dim(\mathbf{E}) = n < \infty$ y T es suprayectiva, entonces $\dim(\mathbf{F}) \leq n$.*
4. *Si $\dim(\mathbf{E}) = \dim(\mathbf{F}) = n < \infty$, T es suprayectiva $\Leftrightarrow T$ es inyectiva.*
5. *Si $\mathbf{E}$ tiene dimensión n y T es biyectiva, entonces*

$$\dim(\mathbf{E}) = \dim(\mathbf{F}).$$

DEMOSTRACIÓN ■ 1. Es evidente.

[8]Recuerde que $E_c(A)$ representa el espacio columna de A (cfr. definición 3.13, pág. 146).

2. $(\Rightarrow)$Si T es suprayectiva, $T(\mathbf{E}) = \mathbf{F}$; y, por tanto, $\dim(T(\mathbf{E})) = \dim(\mathbf{F})$. $(\Leftarrow)$ Si $\dim(T(\mathbf{E})) = \dim(\mathbf{F})$, entonces $T(\mathbf{E}) = \mathbf{F}$.

3. Sea $\{\vec{e}_i\}$, $i = 1, 2, \ldots, n$ una base de $\mathbf{E}$, entonces

$$\mathbf{F} = T(\mathbf{E}) = \text{gn}(T(\vec{e}_1), \ldots, T(\vec{e}_n));$$

luego

$$\dim(\mathbf{F}) \leq n.$$

4. Es consecuencia inmediata de la igualdad (5.2).

5. $\mathbf{F} = T(\mathbf{E})$ por ser T suprayectiva, $\dim(\text{Ker}(T)) = 0$ por ser T inyectiva, y por (5.2) se tiene

$$\dim(\mathbf{F}) = \dim(T(\mathbf{E})) + \dim(\text{Ker}(T)) = n = \dim(\mathbf{E}). \quad \blacksquare$$

Si $T : \mathbf{E} \to \mathbf{F}$ es una transformación lineal biyectiva, entonces existe la transformación inversa de T, $T^{-1} : \mathbf{F} \to \mathbf{E}$; ¿es también lineal T^{-1}? La respuesta es afirmativa.

Teorema 5.11 *Sea $T : \mathbf{E} \to \mathbf{F}$ una transformación lineal biyectiva y sea $T^{-1} : \mathbf{F} \to \mathbf{E}$ la transformación inversa de T, entonces T^{-1} es lineal.*

DEMOSTRACIÓN ■ Sean $\vec{u}_1, \vec{u}_2 \in \mathbf{F}$ y $\alpha, \beta \in \mathbb{R}$. Puesto que T es biyectiva, existen $\vec{u}, \vec{v} \in \mathbf{E}$ tales que $T(\vec{u}) = \vec{u}_1$ y $T(\vec{v}) = \vec{u}_2$; por tanto $T^{-1}(\vec{u}_1) = \vec{u}$ y $T^{-1}(\vec{u}_2) = \vec{v}$. Entonces

$$\begin{aligned} T^{-1}(\alpha\vec{u}_1 + \beta\vec{u}_2) &= T^{-1}(\alpha T(\vec{u}) + \beta T(\vec{v})) \\ &= T^{-1}(T(\alpha\vec{u} + \beta\vec{v})) \\ &= \alpha\vec{u} + \beta\vec{v} \\ &= \alpha T^{-1}(\vec{u}_1) + \beta T^{-1}(\vec{u}_2). \quad \blacksquare \end{aligned}$$

5.2 Representaciones matriciales de transformaciones lineales

En el teorema 5.3 (cfr. pág. 421) vimos que para toda transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ existe una matriz $A \in \mathfrak{M}_{m\times n}$ tal que $T = T_A$; esto es, $T(\vec{u}) = A\vec{u}$ $\forall \vec{u} \in \mathbb{R}^n$. En esta sección veremos que este resultado se puede extender a transformaciones $T \in \mathscr{L}(\mathbf{E}, \mathbf{F})$ cuando los espacios vectoriales $\mathbf{E}$ y $\mathbf{F}$ tienen dimensiones finitas.

5.2.1 Vectores de coordenadas, cambio de bases

Recordemos (cfr. pág. 335) que si $\mathscr{B} = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ es una base fija de un espacio vectorial $\mathbf{E}$, entonces, para cada $\vec{u} \in \mathbf{E}$, existen escalares únicos $\alpha_1, \alpha_2, \ldots, \alpha_n$ tales que

$$\vec{u} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \cdots + \alpha_n\vec{e}_n \tag{5.3}$$

Si convenimos que todo vector se va a escribir respetando el mismo orden de los términos en el lado derecho de (5.3); es decir, un escalar que multiplica a $\vec{e}_1$, más un escalar que multiplica a $\vec{e}_2$, etc., más un escalar que multiplica a $\vec{e}_n$, diremos que $\mathscr{B}$ es una **base ordenada** del espacio $\mathbf{E}$ y escribiremos, con frecuencia, $(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n)$ en lugar de $\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\}$ para hacer énfasis en este hecho.[9] Es claro que si se conoce el vector $(\alpha_1,\alpha_2,\ldots,\alpha_n)$ se conoce el vector $\vec{u}$ y viceversa. Esto motiva la siguiente definición.

Definición 5.10 *Sea $\mathscr{B} = (\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n)$ una base ordenada del espacio $\mathbf{E}$ y $\vec{u} \in \mathbf{E}$. Al único vector $(\alpha_1,\alpha_2,\ldots,\alpha_n) \in \mathbb{R}^n$ tal que*

$$\vec{u} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \cdots + \alpha_n\vec{e}_n$$

*se le llama el **vector coordenado** o el **vector de coordenadas** de $\vec{u}$ **relativo a la base** $\mathscr{B}$, y se denota por[10] $[\vec{u}]_{\mathscr{B}}$; i.e.,*

$$[\vec{u}]_{\mathscr{B}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

► **Ejemplo 5.26** En $\mathbf{P}_3$ sea la base ordenada $\mathscr{B} = (1,x,x^2,x^3)$. Entonces

$$3x-4 = -4(1)+3(x)+0\cdot x^2+0\cdot x^3;$$

luego,

$$[3x-4]_{\mathscr{B}} = \begin{bmatrix} -4 \\ 3 \\ 0 \\ 0 \end{bmatrix}. \blacktriangleleft$$

► **Ejemplo 5.27** En el subespacio

$$S = \text{gn}(1,\cos(2x))$$

de las funciones continuas de $\mathbb{R}$ en $\mathbb{R}$,

$$\cos^2(x) = \frac{1}{2}+\frac{1}{2}\cos(2x).$$

Entonces,

$$\left[\cos^2(x)\right]_{\mathscr{B}} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

para la base ordenada $\mathscr{B} = (1,\cos(2x))$ de S. ◄

[9] De una manera más rigurosa, pero más simple, una *base ordenada* de un espacio vectorial de dimensión finita es una *sucesión finita* de vectores L.I. que genera a este espacio.

[10] Recordemos que hemos convenido en representar un vector de $\mathbb{R}^n$ indistintamente por una n-ada ordenada o por una matriz columna.

▶ **Ejemplo 5.28** Para la base ordenada

$$\mathscr{B}=\left\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right\}$$

de $\mathfrak{M}_{2\times 2}$,

$$\begin{bmatrix}-1&1\\0&2\end{bmatrix}=(-1)\begin{bmatrix}1&0\\0&0\end{bmatrix}+(1)\begin{bmatrix}0&1\\0&0\end{bmatrix}+(0)\begin{bmatrix}0&0\\1&0\end{bmatrix}+(2)\begin{bmatrix}0&0\\0&1\end{bmatrix}.$$

Luego

$$\left[\begin{bmatrix}-1&1\\0&2\end{bmatrix}\right]_{\mathscr{B}}=\begin{bmatrix}-1\\1\\0\\2\end{bmatrix}.$$ ◀

▶ **Ejemplo 5.29** Dado que

$$\begin{bmatrix}1&-1&3\\1&2&-1\\0&-1&2\end{bmatrix}\sim\begin{bmatrix}1&-1&3\\0&3&-4\\0&-1&2\end{bmatrix}\sim\begin{bmatrix}1&-1&3\\0&3&-4\\0&0&2\end{bmatrix},$$

se sigue que $\{(1,1,0),(-1,2,-1),(3,-1,2)\}$ es una base de $\mathbb{R}^3$ (cfr. teorema 3.13, pág. 158). Sea el vector $(2,1,3)$ de $\mathbb{R}^3$ y resolvamos el sistema

$$\begin{bmatrix}1&-1&3\\1&2&-1\\0&-1&2\end{bmatrix}\begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{bmatrix}=\begin{bmatrix}2\\1\\3\end{bmatrix}$$

por el método de Gauss:

$$\left[\begin{array}{rrr|r}1&-1&3&2\\1&2&-1&1\\0&-1&2&3\end{array}\right]\sim\left[\begin{array}{rrr|r}1&-1&3&2\\0&3&-4&-1\\0&-1&2&3\end{array}\right]\sim\left[\begin{array}{rrr|r}1&-1&3&2\\0&3&-4&-1\\0&0&2&8\end{array}\right]$$

$$\therefore\quad\begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{bmatrix}=\begin{bmatrix}-5\\5\\4\end{bmatrix}.$$

Por lo que

$$[(2,1,3)]_{\mathscr{B}}=\begin{bmatrix}-5\\5\\4\end{bmatrix}$$

para la base ordenada $((1,1,0),(-1,2,-1),(3,-1,2))$.

Comprobación:

$$-5(1,1,0)+5(-1,2,-1)+4(3,-1,2)=(2,1,3).◀$$

Las siguientes propiedades de los vectores de coordenadas son sencillas de probar y se dejan como ejercicio al lector.

Teorema 5.12 *Sean* **E** *un espacio vectorial de dimensión n y* $\mathscr{B}$ *una base ordenada de él. Entonces:*

1. $[\vec{u}+\vec{v}]_{\mathscr{B}}=[\vec{u}]_{\mathscr{B}}+[\vec{v}]_{\mathscr{B}} \quad \forall \vec{u},\vec{v}\in\mathbf{E}$.
2. $[\alpha\vec{u}]_{\mathscr{B}}=\alpha[\vec{u}]_{\mathscr{B}} \quad \forall \vec{u}\in\mathbf{E}, \quad \forall\alpha\in\mathbb{R}$.
3. $[\vec{u}]_{\mathscr{B}}=[\vec{v}]_{\mathscr{B}}\Leftrightarrow\vec{u}=\vec{v}$.
4. $[\vec{u}]_{\mathscr{B}}=\vec{0}_{\mathbb{R}^n}\Leftrightarrow\vec{u}=\vec{0}_{\mathbf{E}}$.
5. $\vec{u}_1,\ldots,\vec{u}_k$ *son L.I. en* **E** $\Leftrightarrow [\vec{u}_1]_{\mathscr{B}},\ldots,[\vec{u}_k]_{\mathscr{B}}$ *son L.I. en* $\mathbb{R}^n$.
6. *Para todo* $\vec{x}=(x_1,\ldots,x_n)\in\mathbb{R}^n$ *existe un único* $\vec{u}\in\mathbf{E}$ *tal que* $[\vec{u}]_{\mathscr{B}}=\vec{x}$ *y viceversa.*

En la ulterior discusión necesitaremos del siguiente lema.

Lema 5.1

1. *Sea* $A\in\mathfrak{M}_{m\times n}$ *una matriz tal que* $A\vec{x}=\vec{0}_{\mathbb{R}^m}$ *para todo* $\vec{x}\in\mathbb{R}^n$. *Entonces* $A=\mathscr{O}$, *la matriz cero de tamaño* $m\times n$.
2. *Sea* $A\in\mathfrak{M}_{n\times n}$ *una matriz tal que* $A\vec{x}=\vec{x}$ *para todo* $\vec{x}\in\mathbb{R}^n$. *Entonces* $A=I_n$, *la matriz identidad de orden n.*

DEMOSTRACIÓN ■ 1. Sean $\vec{e}_i$ los vectores de la base canónica de $\mathbb{R}^n$ (la componente i de $\vec{e}_i$ es 1 y las demás son 0). Entonces, si $\vec{K}_i$ es la columna i de A, se tiene

$$\vec{0}_{\mathbb{R}^m}=A\vec{e}_i=\vec{K}_i;$$

de donde $A=\mathscr{O}$.

2. Es consecuencia inmediata del inciso anterior, pues

$$A\vec{x}=\vec{x}\Rightarrow(A-I_n)\vec{x}=\vec{0}_{\mathbb{R}^n}. \quad ■$$

Supongamos ahora que $\mathscr{B}_1=(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n)$ y $\mathscr{B}_2=(\vec{f}_1,\vec{f}_2,\ldots,\vec{f}_n)$ son bases ordenadas de un mismo espacio vectorial **E**. Sea

$$[\vec{u}]_{\mathscr{B}_2}=\begin{bmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_n\end{bmatrix}$$

el vector de coordenadas de $\vec{u}$ relativo a la base $\mathscr{B}_2$. Escribamos cada vector $\vec{f}_i$ como combinación lineal de los elementos de la base ordenada $\mathscr{B}_1$:

$$
\begin{aligned}
\vec{f}_1 &= a_{11}\vec{e}_1 + a_{12}\vec{e}_2 + \cdots + a_{1n}\vec{e}_n,\\
\vec{f}_2 &= a_{21}\vec{e}_1 + a_{22}\vec{e}_2 + \cdots + a_{2n}\vec{e}_n,\\
\vdots &\ \vdots \qquad\qquad \vdots\\
\vec{f}_n &= a_{n1}\vec{e}_1 + a_{n2}\vec{e}_2 + \cdots + a_{nn}\vec{e}_n.
\end{aligned}
$$

Entonces

$$
\begin{aligned}
\vec{u} &= \alpha_1\vec{f}_1 + \alpha_2\vec{f}_2 + \cdots + \alpha_n\vec{f}_n\\
&= \alpha_1\left(a_{11}\vec{e}_1 + a_{12}\vec{e}_2 + \cdots + a_{1n}\vec{e}_n\right)\\
&\quad + \alpha_2\left(a_{21}\vec{e}_1 + a_{22}\vec{e}_2 + \cdots + a_{2n}\vec{e}_n\right)\\
&\quad + \cdots\\
&\quad + \alpha_n\left(a_{n1}\vec{e}_1 + a_{n2}\vec{e}_2 + \cdots + a_{nn}\vec{e}_n\right)\\
&= \left(\alpha_1 a_{11} + \alpha_2 a_{21} + \cdots + \alpha_n a_{n1}\right)\vec{e}_1\\
&\quad + \left(\alpha_1 a_{12} + \alpha_2 a_{22} + \cdots + \alpha_n a_{n2}\right)\vec{e}_2\\
&\quad + \cdots\\
&\quad + \left(\alpha_1 a_{1n} + \alpha_2 a_{2n} + \cdots + \alpha_n a_{nn}\right)\vec{e}_n.
\end{aligned}
$$

Esto es,

$$
\begin{aligned}
[\vec{u}]_{\mathscr{B}_1} &= \begin{bmatrix} \alpha_1 a_{11} + \alpha_2 a_{21} + \cdots + \alpha_n a_{n1}\\ \alpha_1 a_{12} + \alpha_2 a_{22} + \cdots + \alpha_n a_{n2}\\ \vdots\\ \alpha_1 a_{1n} + \alpha_2 a_{2n} + \cdots + \alpha_n a_{nn} \end{bmatrix}\\
&= \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1}\\ a_{12} & a_{22} & \cdots & a_{n2}\\ \vdots & \vdots & \ddots & \vdots\\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n \end{bmatrix}.
\end{aligned}
$$

Entonces, si

$$
P = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1}\\ a_{12} & a_{22} & \cdots & a_{n2}\\ \vdots & \vdots & \ddots & \vdots\\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}
$$

y conocemos el vector de coordenadas del vector $\vec{u}$ relativo a la base $\mathscr{B}_2$, podemos conocer también el vector de coordenadas de $\vec{u}$ relativo a la base $\mathscr{B}_1$ mediante la matriz P:

$$
[\vec{u}]_{\mathscr{B}_1} = P\,[\vec{u}]_{\mathscr{B}_2} \tag{5.4}
$$

Ahora escribamos los vectores de la base ordenada $\mathscr{B}_1$ como combinaciones lineales de los vectores de la base $\mathscr{B}_2$:

$$
\begin{aligned}
\vec{e}_1 &= b_{11}\vec{f}_1 + b_{12}\vec{f}_2 + \cdots + b_{1n}\vec{f}_n,\\
\vec{e}_2 &= b_{21}\vec{f}_1 + b_{22}\vec{f}_2 + \cdots + b_{2n}\vec{f}_n,\\
\vdots &\ \vdots \qquad\qquad \vdots\\
\vec{e}_n &= b_{n1}\vec{f}_1 + b_{n2}\vec{f}_2 + \cdots + b_{nn}\vec{f}_n.
\end{aligned}
$$

Sea

$$Q = \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{n1} \\ b_{12} & b_{22} & \cdots & b_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{1n} & b_{2n} & \cdots & b_{nn} \end{bmatrix},$$

entonces, por analogía, se tiene

$$[\vec{u}]_{\mathscr{B}_2} = Q\,[\vec{u}]_{\mathscr{B}_1}. \tag{5.5}$$

Es decir, podemos calcular el vector de coordenadas de $\vec{u}$ relativo a la base $\mathscr{B}_2$ si conocemos el vector de coordenadas del mismo vector relativo a la base $\mathscr{B}_1$ multiplicando éste por la matriz Q. Ahora bien, de (5.4) y (5.5) tenemos

$$\begin{aligned} [\vec{u}]_{\mathscr{B}_1} &= P\,[\vec{u}]_{\mathscr{B}_2} \\ &= P\left(Q\,[\vec{u}]_{\mathscr{B}_1}\right) \\ &= (PQ)\,[\vec{u}]_{\mathscr{B}_1} \end{aligned}$$

para todo vector $[\vec{u}]_{\mathscr{B}_1} \in \mathbb{R}^n$. Esto implica, por el lema 5.1, que $PQ = I_n$; es decir, P es invertible y $P^{-1} = Q$. Hemos probado así el siguiente teorema.

Teorema 5.13 *Sean $\mathscr{B}_1 = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ y $\mathscr{B}_2 = (\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_n)$ dos bases ordenadas de un mismo espacio vectorial. Sean $a_{ij} \in \mathbb{R}$ escalares tales que*

$$\begin{array}{ccc} \vec{f}_1 & = & a_{11}\vec{e}_1 + a_{12}\vec{e}_2 + \cdots + a_{1n}\vec{e}_n \\ \vec{f}_2 & = & a_{21}\vec{e}_1 + a_{22}\vec{e}_2 + \cdots + a_{2n}\vec{e}_n \\ \vdots & \vdots & \vdots \\ \vec{f}_n & = & a_{n1}\vec{e}_1 + a_{n2}\vec{e}_2 + \cdots + a_{nn}\vec{e}_n \end{array} \tag{5.6}$$

y sea[11]

$$P = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}.$$

Entonces, la matriz P es invertible y para todo $\vec{u} \in \mathbf{E}$ se tiene:

1. $$[\vec{u}]_{\mathscr{B}_1} = P\,[\vec{u}]_{\mathscr{B}_2}.$$
2. $$[\vec{u}]_{\mathscr{B}_2} = P^{-1}\,[\vec{u}]_{\mathscr{B}_1}.$$

Definición 5.11 *Sean las condiciones del teorema anterior.*

1. *A la matriz P se le dice **matriz cambio de base** de la base $\mathscr{B}_2$ a la base $\mathscr{B}_1$.*
2. *A la matriz P^{-1} se le dice **matriz cambio de base** de la base $\mathscr{B}_1$ a la base $\mathscr{B}_2$.*

[11] Observe que P es la transpuesta de la matriz de los coeficientes de los vectores $\vec{e}_i$ del sistema (5.6).

▶ **Ejemplo 5.30** Sean $\vec{e}_1 = (1,0)$, $\vec{e}_2 = (0,1)$, $\vec{f}_1 = (1,1)$, $\vec{f}_2 = (-1,0)$ y $\mathscr{B}_1 = \{\vec{e}_1, \vec{e}_2\}$, $\mathscr{B}_2 = \{\vec{f}_1, \vec{f}_2\}$. Ambas son bases de $\mathbb{R}^2$ y

$$\vec{f}_1 = 1 \cdot \vec{e}_1 + 1 \cdot \vec{e}_2$$
$$\vec{f}_2 = (-1)\vec{e}_1 + 0 \cdot \vec{e}_2.$$

Entonces,

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \text{ y } P^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$$

luego,

$$[u]_{\mathscr{B}_2} = P^{-1}[u]_{\mathscr{B}_1}.$$

Así, por ejemplo, podemos escribir $(-4,3)$ como combinación lineal de $\vec{f}_1, \vec{f}_2$, al hallar $[(-4,3)]_{\mathscr{B}_2}$; pues

$$\begin{aligned} [(-4,3)]_{\mathscr{B}_2} &= P^{-1}[(-4,3)]_{\mathscr{B}_1} \\ &= \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 7 \end{bmatrix}; \end{aligned}$$

y por tanto

$$(-4,3) = 3\vec{f}_1 + 7\vec{f}_2,$$

como el lector puede comprobar fácilmente.◀

▶ **Ejemplo 5.31** En el ejemplo 5.29 vimos que

$$\mathscr{B}_2 = \{(1,1,0), (-1,2,-1), (3,-1,2)\}$$

es una base de $\mathbb{R}^3$. Sea $\mathscr{B}_1 = \{(1,0,0), (0,1,0), (0,1,0)\}$ la base canónica de este espacio. Entonces para este caso, claramente

$$P = \begin{bmatrix} 1 & -1 & 3 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Calculemos P^{-1} por el método de Gauss-Jordan:

$$\left[\begin{array}{rrr|rrr} 1 & -1 & 3 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 1 & -1 & 3 & 1 & 0 & 0 \\ 0 & 3 & -4 & -1 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right]$$
$$\sim \left[\begin{array}{rrr|rrr} 1 & -1 & 3 & 1 & 0 & 0 \\ 0 & 3 & -4 & -1 & 1 & 0 \\ 0 & 0 & 2 & -1 & 1 & 3 \end{array}\right]$$
$$\sim \left[\begin{array}{rrr|rrr} 1 & -1 & 3 & 1 & 0 & 0 \\ 0 & 3 & -4 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1/2 & 1/2 & 3/2 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & -1 & 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & -3 & 3 & 6 \\ 0 & 0 & 1 & -1/2 & 1/2 & 3/2 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 5/2 & -3/2 & -9/2 \\ 0 & 1 & 0 & -1 & 1 & 2 \\ 0 & 0 & 1 & -1/2 & 1/2 & 3/2 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3/2 & -1/2 & -5/2 \\ 0 & 1 & 0 & -1 & 1 & 2 \\ 0 & 0 & 1 & -1/2 & 1/2 & 3/2 \end{array}\right]$$

y por tanto

$$P^{-1} = \begin{bmatrix} 3/2 & -1/2 & -5/2 \\ -1 & 1 & 2 \\ -1/2 & 1/2 & 3/2 \end{bmatrix}.$$

Así, si $[\vec{u}]_{\mathscr{B}_1}$ es el vector de coordenadas de $\vec{u}$ relativo a la base canónica $\mathscr{B}_1$, entonces el vector de coordenadas de éste relativo a la base $\mathscr{B}_2$ es

$$[\vec{u}]_{\mathscr{B}_2} = P^{-1}[\vec{u}]_{\mathscr{B}_1}.$$

Por ejemplo, si $\vec{u} = (2,1,3)$, por tanto, $[\vec{u}]_{\mathscr{B}_1} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ y

$$[\vec{u}]_{\mathscr{B}_2} = \begin{bmatrix} 3/2 & -1/2 & -5/2 \\ -1 & 1 & 2 \\ -1/2 & 1/2 & 3/2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -5 \\ 5 \\ 4 \end{bmatrix}.$$

Que es el mismo resultado que obtuvimos en el ejemplo 5.29.◀

❍ **Nota 5.3** Aunque en el ejemplo anterior no hicimos énfasis por medio de notación, hemos considerado las dos bases como bases ordenadas. Esto lo haremos frecuentemente en aras de brevedad; es decir, toda base, a lo largo de lo que resta de este capítulo, se considera una base ordenada aun si esto no se hace notar explícitamente.

► **Ejemplo 5.32** Sean las bases $\mathscr{B}_1 = \{1, x\}$ y $\mathscr{B}_2 = \{4, 2-x\}$ del espacio de polinomios de grado a lo más 1 (note que en ambos casos claramente los respectivos elementos de cada conjunto son L.I. y recuerde que la dimensión de $\mathbf{P}_1$ es 2; razón por la que ambas son, efectivamente, bases de este espacio).

1. Encontrar la matriz cambio de base de la base $\mathscr{B}_1$ a la base $\mathscr{B}_2$.
2. Hallar el vector de coordenadas del polinomio $p(x) = 4 - 5x$ relativo a la base $\mathscr{B}_2$.
3. Escribir el polinomio $p(x)$ como combinación lineal de los elementos de la base $\mathscr{B}_2$.◀

Solución 1. Tenemos que

$$\begin{aligned} 4 &= 4\cdot 1 + 0\cdot x \\ 2 - x &= 2\cdot 1 + (-1)x \end{aligned}$$

por lo que

$$P = \begin{bmatrix} 4 & 2 \\ 0 & -1 \end{bmatrix}$$

y, por el método de la adjunta,

$$P^{-1} = -\frac{1}{4}\begin{bmatrix} -1 & -2 \\ 0 & 4 \end{bmatrix}.$$

La cual es la matriz cambio de base de $\mathscr{B}_1$ a $\mathscr{B}_2$.

2. Ya que $p(x) = 4 \cdot 1 + (-5)x$,

$$[p(x)]_{\mathscr{B}_1} = \begin{bmatrix} 4 \\ -5 \end{bmatrix}.$$

Entonces

$$\begin{aligned} [p(x)]_{\mathscr{B}_2} &= P^{-1}[p(x)]_{\mathscr{B}_1} \\ &= -\frac{1}{4}\begin{bmatrix} -1 & -2 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 4 \\ -5 \end{bmatrix} \\ &= \begin{bmatrix} -3/2 \\ 5 \end{bmatrix}. \end{aligned}$$

3. Finalmente

$$p(x) = \left(-\frac{3}{2}\right)4 + (5)(2-x)$$

como el lector puede fácilmente verificar. ✔

5.2.2 Representaciones matriciales de un operador lineal

Ahora generalizaremos el concepto de representación matricial que se dio para transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$ a operadores lineales en un espacio vectorial; esto es, para transformaciones lineales de un espacio en sí mismo cuando éste tiene dimensión finita. Para ello, sean $T \in \mathscr{L}(\mathbf{E},\mathbf{E})$ un operador lineal y $\mathscr{B} = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ una base del espacio $\mathbf{E}$. Sean $a_{ij} \in \mathbb{R}$ escalares tales que

$$\begin{aligned} T(\vec{e}_1) &= a_{11}\vec{e}_1 + a_{12}\vec{e}_2 + \cdots + a_{1n}\vec{e}_n \\ T(\vec{e}_2) &= a_{21}\vec{e}_1 + a_{22}\vec{e}_2 + \cdots + a_{2n}\vec{e}_n \\ &\vdots \\ T(\vec{e}_n) &= a_{n1}\vec{e}_1 + a_{n2}\vec{e}_2 + \cdots + a_{nn}\vec{e}_n. \end{aligned} \tag{5.7}$$

Sea $\vec{u} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \cdots + \alpha_n\vec{e}_n \in \mathbf{E}$; entonces $[\vec{u}]_{\mathscr{B}} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ y

$$\begin{aligned} T(\vec{u}) &= T(\alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \cdots + \alpha_n\vec{e}_n) \\ &= \alpha_1 T(\vec{e}_1) + \alpha_2 T(\vec{e}_2) + \cdots + \alpha_n T(\vec{e}_n). \end{aligned}$$

Por tanto

$$\begin{aligned} T(\vec{u}) = \alpha_1&(a_{11}\vec{e}_1 + a_{12}\vec{e}_2 + \cdots + a_{1n}\vec{e}_n) \\ &+\alpha_2(a_{21}\vec{e}_1 + a_{22}\vec{e}_2 + \cdots + a_{2n}\vec{e}_n) \\ &+\cdots \\ &+\alpha_n(a_{n1}\vec{e}_1 + a_{n2}\vec{e}_2 + \cdots + a_{nn}\vec{e}_n) \end{aligned}$$

$$
\begin{aligned}
&= (\alpha_1 a_{11} + \alpha_2 a_{21} + \cdots + \alpha_n a_{n1})\vec{e}_1 \\
&\quad + (\alpha_1 a_{12} + \alpha_2 a_{22} + \cdots + \alpha_n a_{n2})\vec{e}_2 \\
&\quad + \cdots \\
&\quad + (\alpha_1 a_{1n} + \alpha_2 a_{2n} + \cdots + \alpha_n a_{nn})\vec{e}_n;
\end{aligned}
$$

luego

$$
\begin{aligned}
[T(\vec{u})]_{\mathscr{B}} &= \begin{bmatrix} \alpha_1 a_{11} + \alpha_2 a_{21} + \cdots + \alpha_n a_{n1} \\ \alpha_1 a_{12} + \alpha_2 a_{22} + \cdots + \alpha_n a_{n2} \\ \vdots \\ \alpha_1 a_{1n} + \alpha_2 a_{2n} + \cdots + \alpha_n a_{nn} \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.
\end{aligned}
$$

Lo cual motiva la definición 5.12 y prueba el teorema 5.14.

Definición 5.12 *Sean $T : \mathbf{E} \to \mathbf{E}$ un operador lineal en el espacio $\mathbf{E}$, $\mathscr{B} = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ una base (ordenada) de este espacio y $a_{ij} \in \mathbb{R}$ escalares tales que se tiene el sistema de igualdades (5.7); a la matriz denotada y definida como*

$$
[T]_{\mathscr{B}} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}
$$

se[12] *le llama la* ***representación matricial*** *del operador T relativa a la base $\mathscr{B}$.*

Teorema 5.14 *Sean $\mathbf{E}$, T, $\mathscr{B}$ y $[T]_{\mathscr{B}}$ como en la definición 5.12. Entonces*

$$
[T(\vec{u})]_{\mathscr{B}} = [T]_{\mathscr{B}} [\vec{u}]_{\mathscr{B}} \tag{5.8}
$$

para todo $\vec{u} \in \mathbf{E}$.

► **Ejemplo 5.33** En $\mathbf{P}_3$ sean $T : \mathbf{P}_3 \to \mathbf{P}_3$ el operador lineal definido por $T(p) = p'$, $\mathscr{B} = \{1, x, x^2, x^3\}$, y $p(x) = -3 + 8x - 5x^2 - 7x^3 \in \mathbf{P}_3$. Entonces:

$$
\begin{aligned}
T(1) &= 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\
T(x) &= 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\
T(x^2) &= 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\
T(x^3) &= 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2 + 0 \cdot x^3.
\end{aligned}
$$

[12]Observe que $[T]_{\mathscr{B}}$ es la matriz transpuesta de coeficientes de los $\vec{e}_i$ del sistema de igualdades (5.7).

Así,

$$[T]_{\mathscr{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

y

$$T(p(x)) = 8 - 10x - 21x^2 = 8 \cdot 1 + (-10) \cdot x + (-21) \cdot x^2 + 0 \cdot x^3.$$

Por tanto,

$$[T(p(x))]_{\mathscr{B}} = \begin{bmatrix} 8 \\ -10 \\ -21 \\ 0 \end{bmatrix}.$$

Por otro lado,

$$[T]_{\mathscr{B}}[p(x)]_{\mathscr{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -3 \\ 8 \\ -5 \\ -7 \end{bmatrix} = \begin{bmatrix} 8 \\ -10 \\ -21 \\ 0 \end{bmatrix} = [T(p(x))]_{\mathscr{B}}.$$ ◀

❍ **Nota 5.4** Supongamos que T es un operador lineal en un espacio de dimensión finita $\mathbf{E}$; con $[T]_{\mathscr{B}}$ la representación de este operador relativa a una base $\mathscr{B}$. Supongamos que A es una matriz cuadrada de orden n tal que

$$[T(\vec{u})]_{\mathscr{B}} = A\,[\vec{u}]_{\mathscr{B}} \tag{5.9}$$

para todo $\vec{u} \in \mathbf{E}$. Entonces

$$[T]_{\mathscr{B}}\,[\vec{u}]_{\mathscr{B}} = [T(\vec{u})]_{\mathscr{B}} = A\,[\vec{u}]_{\mathscr{B}};$$

y por tanto

$$([T]_{\mathscr{B}} - A)\vec{x} = \vec{0}_{\mathbb{R}^n} \quad \forall \vec{x} \in \mathbb{R}^n.$$

Por el lema 5.1 se tiene entonces $[T]_{\mathscr{B}} - A = \mathscr{O}$; esto es,

$$[T]_{\mathscr{B}} = A.$$

Es decir, la única matriz cuadrada de orden n que satisface (5.9) para todo $\vec{u} \in \mathbf{E}$ es la representación matricial, $[T]_{\mathscr{B}}$, del operador T relativa a la base $\mathscr{B}$.

Como acabamos de ver, en la nota 5.4, la representación matricial de un operador lineal, relativa a una base $\mathscr{B}$, es la única matriz que satisface (5.9); sin embargo, si se toma otra base $\mathscr{B}_1$ del espacio $\mathbf{E}$, también se cumple (5.8) del teorema 5.14 con $\mathscr{B} = \mathscr{B}_1$. ¿Cómo están relacionadas entonces las representaciones matriciales de un mismo operador lineal relativas a distintas bases? Para responder esta pregunta supongamos que $\mathscr{B}_1$ y $\mathscr{B}_2$ son distintas bases del mismo espacio vectorial $\mathbf{E}$ y $[T]_{\mathscr{B}_1}$, $[T]_{\mathscr{B}_2}$ son sendas representaciones matriciales del operador T relativas a estas bases. Sea P la matriz

cambio de base de la base $\mathscr{B}_2$ a la base $\mathscr{B}_1$; entonces $[\vec{u}]_{\mathscr{B}_2} = P^{-1}\,[\vec{u}]_{\mathscr{B}_1}$ y $[\vec{u}]_{\mathscr{B}_1} = P\,[\vec{u}]_{\mathscr{B}_2}$ para todo $\vec{u} \in \mathbf{E}$. Luego

$$\begin{aligned}
[T(\vec{u})]_{\mathscr{B}_2} &= P^{-1}\,[T(\vec{u})]_{\mathscr{B}_1}\\
&= P^{-1}\left([T]_{\mathscr{B}_1}\,[\vec{u}]_{\mathscr{B}_1}\right)\\
&= \left(P^{-1}\,[T]_{\mathscr{B}_1}\right)\left(P\,[\vec{u}]_{\mathscr{B}_2}\right)\\
&= \left(P^{-1}\,[T]_{\mathscr{B}_1}\,P\right)[\vec{u}]_{\mathscr{B}_2}
\end{aligned}$$

para todo $\vec{u} \in \mathbf{E}$. Por la unicidad de la representación matricial de un operador mostrada en la nota 5.4, se tiene que

$$[T]_{\mathscr{B}_2} = P^{-1}\,[T]_{\mathscr{B}_1}\,P.$$

Hemos probado así el siguiente teorema.

Teorema 5.15 *Sean $\mathbf{E}$ un espacio vectorial; $\mathscr{B}_1$ y $\mathscr{B}_2$ bases del mismo; T un operador lineal en $\mathbf{E}$; y $[T]_{\mathscr{B}_1}$, $[T]_{\mathscr{B}_2}$ las respectivas representaciones matriciales de T relativas a estas bases. Si P es la matriz cambio de base de $\mathscr{B}_2$ a $\mathscr{B}_1$, entonces*

$$[T]_{\mathscr{B}_2} = P^{-1}\,[T]_{\mathscr{B}_1}\,P.$$

► **Ejemplo 5.34** Sea el operador lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$, definido por $T(x,y) = (4x-2y, 2x+y)$ y $\mathscr{B}_1 = \{(1,0)(0,1)\}$, $\mathscr{B}_2 = \{(1,1),(1,0)\}$. Entonces

$$\begin{aligned}
T(1,0) &= (4,2) &&= 4(1,0)+2(0,1),\\
T(0,1) &= (-2,1) &&= -2(1,0)+1(0,1),
\end{aligned}$$

y

$$\begin{aligned}
(1,1) &= 1(1,0)+1(1,0),\\
(1,0) &= 1(1,0)+0(0,1).
\end{aligned}$$

Así que

$$[T]_{\mathscr{B}_1} = \begin{bmatrix} 4 & -2\\ 2 & 1 \end{bmatrix}$$

y

$$P = \begin{bmatrix} 1 & 1\\ 1 & 0 \end{bmatrix},\ P^{-1} = \begin{bmatrix} 0 & 1\\ 1 & -1 \end{bmatrix}.$$

Luego,

$$\begin{aligned}
[T]_{\mathscr{B}_2} &= P^{-1}[T]_{\mathscr{B}_1}P\\
&= \begin{bmatrix} 0 & 1\\ 1 & -1 \end{bmatrix}\begin{bmatrix} 4 & -2\\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1\\ 1 & 0 \end{bmatrix}\\
&= \begin{bmatrix} 2 & 1\\ 2 & -3 \end{bmatrix}\begin{bmatrix} 1 & 1\\ 1 & 0 \end{bmatrix}\\
&= \begin{bmatrix} 3 & 2\\ -1 & 2 \end{bmatrix}. \blacktriangleleft
\end{aligned}$$

Definición 5.13 *Si A y B son matrices cuadradas de orden n, se dice que A es* ***similar*** *a B si existe una matriz cuadrada C del mismo orden tal que*

$$B = C^{-1}AC.$$

Para denotar que la matriz A es similar a la matriz B escribiremos $A \simeq B$.

Las siguientes propiedades, contenidas en el teorema 5.16, son sencillas de probar y la demostración de cada una de ellas se deja como ejercicio al lector.

Teorema 5.16 *Sean $A, B, C, D \in \mathfrak{M}_n$. Entonces:*

1. $A \simeq A$.
2. $A \simeq B \Rightarrow B \simeq A$.
3. $A \simeq B$ y $B \simeq D \Rightarrow A \simeq D$.
4. *Si* $A \simeq B$, *entonces* $\det(A) = \det(B)$.
5. *Si* $A \simeq B$, *entonces A es invertible si y sólo si B es invertible.*

❍ **Nota 5.5**

1. Por el inciso 2 del teorema precedente, $A \simeq B \Rightarrow B \simeq A$, diremos que dos matrices son similares si cumplen con la definición 5.13.
2. El teorema 5.15 implica que dos representaciones cualesquiera de un operador lineal $T \in \mathscr{L}(\mathbf{E}, \mathbf{E})$ son similares.

Por el teorema 5.16, inciso 4, cualquier par de representaciones matriciales relativas a diferentes bases de un operador lineal en un espacio tienen el mismo determinante, pues son matrices similares. Hacemos patente este hecho en la siguiente definición.

Definición 5.14 ***(Determinante de un operador)*** *Si* $\mathbf{E}$ *es un espacio de dimensión finita y T es un operador lineal en este espacio, se define el determinante de T,* $\det(T)$*, como el determinante de cualquier representación matricial de T.*

► **Ejemplo 5.35** Sea T el operador lineal definido en el ejemplo 5.34. Vimos en este ejemplo que si $\mathscr{B}_1 = \{(1,0)(0,1)\}$ y $\mathscr{B}_2 = \{(1,1),(1,0)\}$,

$$[T]_{\mathscr{B}_1} = \begin{bmatrix} 4 & -2 \\ 2 & 1 \end{bmatrix} \text{ y } [T]_{\mathscr{B}_2} = \begin{bmatrix} 3 & 2 \\ -1 & 2 \end{bmatrix},$$

entonces

$$\det(T) = \begin{vmatrix} 4 & -2 \\ 2 & 1 \end{vmatrix} = \begin{vmatrix} 3 & 2 \\ -1 & 2 \end{vmatrix} = 8. \blacktriangleleft$$

❍ **Nota 5.6** Dado que dos representaciones matriciales cualesquiera de un operador son similares, por el teorema 5.16, si una de ellas es una matriz invertible todas las demás también lo son.

Vimos en el teorema 5.9, inciso 5 (cfr. pág. 432), que para que un operador lineal definido en $\mathbb{R}^n$ sea inyectivo (y por tanto biyectivo) es necesario y suficiente que su representación matricial (relativa a la base canónica) sea una matriz invertible. Tenemos el caso general de operadores lineales en espacios vectoriales contemplado en el siguiente teorema.

Teorema 5.17 *Sea* $\mathbf{E}$ *un espacio vectorial de dimensión* n *y* $T:\mathbf{E}\to\mathbf{E}$ *un operador lineal. Entonces* T *es inyectivo (y por tanto biyectivo) si y sólo si cualquier representación matricial de* T *es una matriz invertible (si y sólo si* $\det(T)\neq 0$*); en tal caso, si* $[T]_{\mathscr{B}}$ *es la representación matricial de* T *relativa a una base* $\mathscr{B}$*,* $[T]_{\mathscr{B}}^{-1}$ *es la representación matricial del operador lineal inverso*[13] T^{-1} *relativa a la misma base.*

DEMOSTRACIÓN ■ Sea $[T]_{\mathscr{B}}$ una representación matricial del operador T.

1. Supongamos que la matriz $[T]_{\mathscr{B}}$ es no singular y sea $\vec{u}\in\operatorname{Ker}(T)$. Entonces

$$[T]_{\mathscr{B}}[\vec{u}]_{\mathscr{B}}=[T(\vec{u})]_{\mathscr{B}}=\vec{0}_{\mathbb{R}^n}.$$

Luego

$$([T]_{\mathscr{B}})^{-1}([T]_{\mathscr{B}}[\vec{u}]_{\mathscr{B}})=([T]_{\mathscr{B}})^{-1}\vec{0}_{\mathbb{R}^n}=\vec{0}_{\mathbb{R}^n}$$

y por tanto,

$$[\vec{u}]_{\mathscr{B}}=\vec{0}_{\mathbb{R}^n};$$

lo cual implica

$$\vec{u}=\vec{0}_{\mathbf{E}}.$$

2. Supongamos ahora que T es inyectivo, entonces T es suprayectivo pues

$$\underbrace{\dim(\operatorname{Ker}(T))}_{0}+\dim(T(\mathbf{E}))=n.$$

Por tanto, T es biyectivo. Sea $T^{-1}:\mathbf{E}\to\mathbf{E}$ la transformación inversa de T; sabemos, del teorema 5.11, que T^{-1} es también lineal. Sea $\left[T^{-1}\right]_{\mathscr{B}}$ la representación matricial de T^{-1} relativa a la base $\mathscr{B}$. Sea $\vec{v}\in\mathbf{E}$, como T es suprayectiva existe $\vec{u}\in\mathbf{E}$ tal que $T(\vec{u})=\vec{v}$; i.e., $\vec{u}=T^{-1}(\vec{v})$. Entonces

$$\begin{aligned}\left[T^{-1}(\vec{v})\right]_{\mathscr{B}} &= [T^{-1}]_{\mathscr{B}}[\vec{v}]_{\mathscr{B}}\\ &= \left[T^{-1}\right]_{\mathscr{B}}[T(\vec{u})]_{\mathscr{B}}\\ &= \left[T^{-1}\right]_{\mathscr{B}}[T]_{\mathscr{B}}[\vec{u}]_{\mathscr{B}}\\ &= \left(\left[T^{-1}\right]_{\mathscr{B}}[T]_{\mathscr{B}}\right)[T^{-1}(\vec{v})]_{\mathscr{B}}.\end{aligned}$$

[13] Recuerde que la transformación inversa de un operador lineal también es un operador lineal (cfr. teorema 5.11).

Por tanto $\left(\left[T^{-1}\right]_{\mathscr{B}}[T]_{\mathscr{B}}\right)\vec{x}=\vec{x}$ para todo[14] $\vec{x}\in\mathbb{R}^n$. Del lema 5.1 (cfr. pág. 436) se sigue que $\left[T^{-1}\right]_{\mathscr{B}}[T]_{\mathscr{B}}=I_n$. Luego $[T]_{\mathscr{B}}$ es invertible y $[T]_{\mathscr{B}}^{-1}=\left[T^{-1}\right]_{\mathscr{B}}$. ■

5.2.3 Representaciones matriciales de transformaciones lineales

Así como es posible encontrar representaciones matriciales de operadores lineales, también se pueden hallar representaciones matriciales para transformaciones lineales en general. El procedimiento es completamente análogo al que realizamos en la subsección anterior; por lo que dejamos al lector que llene los detalles, en cuanto a deducciones y demostraciones, en forma similar a como lo hicimos antes, de las definiciones y teoremas que damos a continuación.

Definición 5.15 *Sean* **E**, **F** *espacios de dimensiones finitas;* $\mathscr{B}_1=\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\}$, $\mathscr{B}_2=\{\vec{f}_1,\vec{f}_2,\ldots,\vec{f}_m\}$ *bases de sendos espacios;* $T\in\mathscr{L}(\mathbf{E},\mathbf{F})$ *y* a_{ij} *escalares tales que*

$$\begin{array}{rcl} T(\vec{e}_1) & = & a_{11}\vec{f}_1+a_{12}\vec{f}_2+\cdots+a_{1m}\vec{f}_m \\ T(\vec{e}_2) & = & a_{21}\vec{f}_1+a_{22}\vec{f}_2+\cdots+a_{2m}\vec{f}_m \\ \vdots & \vdots & \vdots \\ T(\vec{e}_n) & = & a_{n1}\vec{f}_1+a_{n2}\vec{f}_2+\cdots+a_{nm}\vec{f}_m \end{array} \tag{5.10}$$

Se define la ***representación matricial*** *de T relativa a las bases* $\mathscr{B}_1$ *y* $\mathscr{B}_2$ *como*[15]

$$[T]_{\mathscr{B}_1}^{\mathscr{B}_2}=\begin{bmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \ldots & a_{nm} \end{bmatrix}.$$

Teorema 5.18 *Sean* **E**, $\mathscr{B}_1$, $\mathscr{B}_2$ *y T como en la definición precedente, entonces*

$$[T(\vec{u})]_{\mathscr{B}_2}=[T]_{\mathscr{B}_1}^{\mathscr{B}_2}[\vec{u}]_{\mathscr{B}_1}\quad\forall\vec{u}\in\mathbf{E}.$$

► **Ejemplo 5.36** Sea la transformación lineal $T:\mathbb{R}^3\to\mathbb{R}^2$ definida por $T(x,y,z)=(3x+2y-4z,x-5y+3z)$ y $\mathscr{B}_1=\{(1,1,1),(1,1,0),(1,0,0)\}$, $\mathscr{B}_2=\{(1,3),(2,5)\}$. Encontrar la representación matricial de T relativa a las bases $\mathscr{B}_1$ y $\mathscr{B}_2$.◄

Solución

$$\begin{array}{rcl} T(1,1,1) & = & (1,-1), \\ T(1,1,0) & = & (5,-4), \\ T(1,0,0) & = & (3,1). \end{array}$$

[14] Es fácil demostrar, lo cual se deja como ejercicio al lector, que si $\vec{x}=(x_1,\ldots,x_n)\in\mathbb{R}^n$, existe $\vec{v}\in\mathbf{E}$ tal que $[\vec{v}]_{\mathscr{B}}=\vec{x}$.

[15] Note que $[T]_{\mathscr{B}_1}^{\mathscr{B}_2}$ es la matriz transpuesta de los coeficientes de los vectores $\vec{f}_i$ del sistema de igualdades (5.10).

Se requiere encontrar escalares a_{ij} tales que

$$\begin{aligned}(1,-1) &= a_{11}(1,3)+a_{12}(2,5),\\(5,-4) &= a_{21}(1,3)+a_{22}(2,5),\\(3,1) &= a_{31}(1,3)+a_{32}(2,5).\end{aligned}$$

Esto es, las soluciones de los sistemas (con la misma matriz de coeficientes):

$$\begin{bmatrix}1 & 2\\3 & 5\end{bmatrix}\begin{bmatrix}a_{11}\\a_{12}\end{bmatrix}=\begin{bmatrix}1\\-1\end{bmatrix},\quad \begin{bmatrix}1 & 2\\3 & 5\end{bmatrix}\begin{bmatrix}a_{21}\\a_{22}\end{bmatrix}=\begin{bmatrix}5\\-4\end{bmatrix}\quad \text{y}$$

$$\begin{bmatrix}1 & 2\\3 & 5\end{bmatrix}\begin{bmatrix}a_{31}\\a_{32}\end{bmatrix}=\begin{bmatrix}3\\1\end{bmatrix}.$$

Resolvamos estos sistemas con la misma matriz de coeficientes (cfr. el método expuesto en la pág. 27).

$$\left[\begin{array}{cc|ccc}1 & 2 & 1 & 5 & 3\\3 & 5 & -1 & -4 & 1\end{array}\right]\sim\left[\begin{array}{cc|ccc}1 & 2 & 1 & 5 & 3\\0 & -1 & -4 & -19 & -8\end{array}\right]$$
$$\sim\left[\begin{array}{cc|ccc}1 & 0 & -7 & -33 & -13\\0 & -1 & -4 & -19 & -8\end{array}\right];$$

$$\therefore\quad \begin{bmatrix}a_{11}\\a_{12}\end{bmatrix}=\begin{bmatrix}-7\\4\end{bmatrix};\quad \begin{bmatrix}a_{21}\\a_{22}\end{bmatrix}=\begin{bmatrix}-33\\19\end{bmatrix}\quad \text{y}$$
$$\begin{bmatrix}a_{31}\\a_{32}\end{bmatrix}=\begin{bmatrix}-13\\8\end{bmatrix}.$$

Entonces,

$$[T]_{\mathscr{B}_1}^{\mathscr{B}_2}=\begin{bmatrix}-7 & -33 & -13\\4 & 19 & 8\end{bmatrix}. \checkmark$$

❍ **Nota 5.7** Sean $\mathscr{B}_1$, $\mathscr{B}_2$ bases de los espacios $\mathbf{E}$ y $\mathbf{F}$, respectivamente. Si $T\in\mathscr{L}(\mathbf{E},\mathbf{F})$ y A es una matriz de tamaño $m\times n$ tal que

$$A[\vec{u}]_{\mathscr{B}_1}=[T(\vec{u})]_{\mathscr{B}_2}\ \forall\vec{u}\in\mathbf{E} \tag{5.11}$$

se tiene entonces

$$(A-[T]_{\mathscr{B}_1}^{\mathscr{B}_2})\vec{x}=\vec{0}_{\mathbb{R}^m}\quad\forall\vec{x}\in\mathbb{R}^n;$$

por el lema 5.1 se sigue que[16]

$$A-[T]_{\mathscr{B}_1}^{\mathscr{B}_2}=\mathscr{O}.$$

Es decir, la única matriz que cumple (5.11) es la representación matricial de T relativa a estas bases del teorema 5.18.

[16] Cfr. nota 5.4, página 443.

Veamos ahora qué sucede con una representación matricial cuando se cambian bases en los espacios vectoriales.

Teorema 5.19 *Sean* **E**, **F** *espacios de dimensiones finitas;* $\mathscr{B}_1$, $\mathscr{B}'_1$ *bases de* **E***;* $\mathscr{B}_2$, $\mathscr{B}'_2$ *bases de* **F***; y* $T \in \mathscr{L}(\mathbf{E}, \mathbf{F})$. *Sean P la matriz cambio de base de* $\mathscr{B}'_1$ *a* $\mathscr{B}_1$ *(i.e.* $[\vec{u}]_{\mathscr{B}_1} = P[\vec{u}]_{\mathscr{B}'_1}$*) y Q la matriz cambio de base de* $\mathscr{B}'_2$ *a* $\mathscr{B}_2$ *(es decir,* $[\vec{v}]_{\mathscr{B}_2} = Q[\vec{v}]_{\mathscr{B}'_2}$*), entonces*

$$[T]_{\mathscr{B}'_1}^{\mathscr{B}'_2} = Q^{-1}\,[T]_{\mathscr{B}_1}^{\mathscr{B}_2}\,P.$$

DEMOSTRACIÓN ■ Sea $\vec{u} \in \mathbf{E}$, entonces

$$\begin{aligned}
(Q^{-1}[T]_{\mathscr{B}_1}^{\mathscr{B}_2}P)[\vec{u}]_{\mathscr{B}'_1} &= (Q^{-1}[T]_{\mathscr{B}_1}^{\mathscr{B}_2})P[\vec{u}]_{\mathscr{B}'_1} \\
&= (Q^{-1}[T]_{\mathscr{B}_1}^{\mathscr{B}_2})[\vec{u}]_{\mathscr{B}_1} \\
&= Q^{-1}([T]_{\mathscr{B}_1}^{\mathscr{B}_2}[\vec{u}]_{\mathscr{B}_1}) \\
&= Q^{-1}[T(\vec{u})]_{\mathscr{B}_2} \\
&= [T(\vec{u})]_{\mathscr{B}'_2}
\end{aligned}$$

y de la nota 5.7 se desprende que

$$Q^{-1}[T]_{\mathscr{B}_1}^{\mathscr{B}_2}P = [T]_{\mathscr{B}'_1}^{\mathscr{B}'_2}. \quad ■$$

► **Ejemplo 5.37** Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ la transformación lineal del ejemplo 5.36; es decir,

$$T(x,y,z) = (3x+2y-4z, x-5y+3z).$$

Sean $\mathscr{B}_1$ la base canónica de $\mathbb{R}^3$ y la base $\mathscr{B}'_1 = \{(1,1,1),(1,1,0),(1,0,0)\}$; $\mathscr{B}_2$ la base canónica de $\mathbb{R}^2$ y la base $\mathscr{B}'_2 = \{(1,3),(2,5)\}$. Entonces, puesto que

$$\begin{aligned}
T(1,0,0) &= (3,1) &&= 3(1,0)+1(0,1) \\
T(0,1,0) &= (2,-5) &&= 2(1,0)-5(0,1) \\
T(0,0,1) &= (-4,3) &&= -4(1,0)+3(0,1)
\end{aligned}$$

se tiene

$$[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} 3 & 2 & -4 \\ 1 & -5 & 3 \end{bmatrix}.$$

Por otra parte, ya que

$$\begin{aligned}
(1,1,1) &= 1(1,0,0)+1(0,1,0)+1(0,0,1) \\
(1,1,0) &= 1(1,0,0)+1(0,1,0)+0(0,0,1) \\
(1,0,0) &= 1(1,0,0)+0(0,1,0)+(0,0,1)
\end{aligned}$$

y

$$\begin{aligned}
(1,3) &= 1(1,0)+3(0,1) \\
(2,5) &= 2(1,0)+5(0,1),
\end{aligned}$$

se concluye que

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, Q = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}, \text{ y } Q^{-1} = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}.$$

Entonces

$$\begin{aligned} [T]_{\mathscr{B}_1'}^{\mathscr{B}_2'} &= Q^{-1}\,[T]_{\mathscr{B}_1}^{\mathscr{B}_2}\,P \\ &= \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 3 & 2 & -4 \\ 1 & -5 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{bmatrix}. \end{aligned}$$

Que es el mismo resultado que obtuvimos en el ejemplo 5.36.◀

Representaciones diagonales de transformaciones lineales

Toda transformación lineal de un espacio en otro tiene diversas representaciones matriciales según las bases con las que se trabaje. De entre todas estas representaciones sería conveniente buscar aquellas que sean suficientemente "sencillas". De la experiencia adquirida a lo largo de este libro, sabemos que las matrices con mayor número de entradas nulas y, más aún, aquellas que son muy parecidas a una matriz diagonal, son matrices que se prestan mucho para minimizar cálculos y conjeturar propiedades de las mismas y, por tanto, de las transformaciones que representan. En el siguiente teorema, dada una transformación lineal, veremos cómo es posible construir bases de tal manera que la correspondiente representación sea una matriz con entradas nulas fuera de la diagonal principal y unos en ella.

Teorema 5.20 *Sean* $\mathbf{E}$ *y* $\mathbf{F}$ *espacios vectoriales de dimensiones* n *y* m*, respectivamente, y* $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$*. Entonces existen bases* $\mathscr{B}_1 = \{e_1,\ldots,\vec{e}_n\}$ *y* $\mathscr{B}_2 = \{f_1,\ldots,f_m\}$*, de sendos espacios, tales que la representación matricial de la transformación lineal relativa a estas bases,* $[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = [a_{ij}]$*, tiene unos en la diagonal principal y ceros fuera de ella; esto es,*

$$a_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{en otro caso} \end{cases} \tag{5.12}$$

DEMOSTRACIÓN ■ Sea r el rango de T; es decir, $\dim(T(\mathbf{E})) = r$. Sea $\{\vec{f}_1,\ldots,\vec{f}_r\}$ una base de $T(\mathbf{E})$ y completemos ésta a una base de $\mathbf{F}$;[17] digamos

$$\mathscr{B}_2 = \{\vec{f}_1,\ldots,\vec{f}_r,\vec{f}_{r+1},\ldots,\vec{f}_m\}.$$

Puesto que $\vec{f}_1,\ldots,\vec{f}_r \in T(\mathbf{E})$, existen $\vec{e}_1,\ldots,\vec{e}_r \in \mathbf{E}$ tales que

$$T(\vec{e}_i) = \vec{f}_i \quad \text{para } i = 1,\ldots,r \tag{5.13}$$

[17] Cfr. el procedimiento que se dio en la página 167 para completar un subconjunto L.I. a una base de un espacio.

Sea $k = \dim(\mathrm{Ker}(T))$, entonces, por (5.2) del teorema 5.8 (cfr. pág. 431), $n = r + k$. Sea $\{\vec{e}_{r+1}, \ldots, \vec{e}_{r+k}\}$ una base de $\mathrm{Ker}(T)$. Por construcción, $\mathscr{B}_2$ es una base de $\mathbf{F}$. Afirmamos que

$$\mathscr{B}_1 = \{\vec{e}_1, \ldots, \vec{e}_r, \vec{e}_{r+1}, \ldots, \vec{e}_{r+k}\}$$

es una base de $\mathbf{E}$. En efecto, puesto que $\mathscr{B}_1$ tiene n elementos, basta probar que éstos son L.I., lo cual es cierto pues si $\alpha_1, \ldots, \alpha_r, \alpha_{r+1}, \ldots, \alpha_{r+k}$ son n-escalares tales que

$$\alpha_1 \vec{e}_1 + \cdots + \alpha_r \vec{e}_r + \alpha_{r+1} \vec{e}_{r+1} + \cdots + \alpha_{r+k} \vec{e}_{r+k} = \vec{0}_{\mathbf{E}} \tag{5.14}$$

Entonces, dado que T es lineal,

$$\alpha_1 T(\vec{e}_1) + \cdots + \alpha_r T(\vec{e}_r) + \alpha_{r+1} T(\vec{e}_{r+1}) + \cdots + \alpha_{r+k} T(\vec{e}_{r+k}) = \vec{0}_{\mathbf{F}}$$

y puesto que $\vec{e}_{r+1}, \ldots, \vec{e}_{r+k} \in \mathrm{Ker}(T)$, de (5.13) y la precedente igualdad se concluye que

$$\alpha_1 \vec{f}_1 + \cdots + \alpha_r \vec{f}_r = \vec{0}_{\mathbf{F}}. \tag{5.15}$$

Como los vectores $\vec{f}_1, \ldots, \vec{f}_r$ son L.I. (pues forman una base de $T(\mathbf{E})$), (5.15) implica

$$\alpha_1 = \cdots = \alpha_r = 0 \tag{5.16}$$

De (5.16) y (5.14) se tiene

$$\alpha_{r+1} \vec{e}_{r+1} + \cdots + \alpha_{r+k} \vec{e}_{r+k} = \vec{0}_{\mathbf{E}} \tag{5.17}$$

y ya que $\vec{e}_{r+1}, \ldots, \vec{e}_{r+k}$ son L.I. (pues forman una base de $\mathrm{Ker}(T)$), (5.17) implica

$$\alpha_{r+1} = \cdots = \alpha_{r+k} = 0 \tag{5.18}$$

De (5.16) y (5.18) se desprende que los vectores $\vec{e}_1, \ldots, \vec{e}_r, \vec{e}_{r+1}, \ldots, \vec{e}_{r+k}$ son L.I. Finalmente, puesto que

$$T(\vec{e}_i) = \vec{f}_i = 0 \cdot \vec{f}_1 + \cdots + 0 \cdot \vec{f}_{i-1} + 1 \cdot \vec{f}_i + 0 \cdot \vec{f}_{i+1} + \cdots + 0 \cdot \vec{f}_m$$

para $i = 1, \ldots, r$ y

$$T(\vec{e}_j) = \vec{0}_{\mathbf{F}} = 0 \cdot \vec{f}_1 + \cdots + 0 \cdot \vec{f}_m$$

para $j = r+1, \ldots, r+k$, se sigue que

$$[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = [a_{ij}]$$

con los a_{ij} dados por (5.12). ■

► **Ejemplo 5.38** Sea $T : \mathbf{P}_3 \to \mathbf{P}_2$ la transformación derivación; esto es, $T(p) = p'$. Construyamos, siguiendo el proceso dado en la demostración del teorema 5.20, bases de $\mathbf{P}_3$ y $\mathbf{P}_2$ tales que la representación matricial de T relativa a éstas tenga unos en la diagonal principal y ceros fuera de ella. Para ello notemos antes los dos siguientes hechos:

1. $T(\mathbf{P}_3) = \mathbf{P}_2$ (T es suprayectiva). Efectivamente, si $q(x) = ax^2 + bx + c \in \mathbf{P}_2$, entonces $p(x) = \frac{1}{3}ax^3 + \frac{1}{2}bx^2 + cx$ satisface

$$T(p(x)) = p'(x) = ax^2 + bx + c = q(x).$$

2. $\operatorname{Ker}(T) = \operatorname{gn}(1)$. En efecto,

$$p \in \operatorname{Ker}(T) \Leftrightarrow p'(x) = 0 \Leftrightarrow p(x) = c \Leftrightarrow p \in \operatorname{gn}(1).$$

Por el inciso 1 podemos tomar como base de $T(\mathbf{P}_3)$ la base canónica de $\mathbf{P}_2$, $\mathscr{B}_2 = \{1, x, x^2\}$ ($f_1 = 1$, $f_2 = x$, $f_3 = x^2$). Sean $e_1 = x$, $e_2 = \frac{1}{2}x^2$ y $e_3 = \frac{1}{3}x^3$, entonces $T(e_1) = 1$, $T(e_2) = x$ y $T(e_3) = x^2$ (los valores de los e_i se encuentran por simple inspección o integrando). Siguiendo el proceso del teorema 5.20, la base $\mathscr{B}_1$ se construye adjuntando a e_1, e_2 y e_3 los elementos de la base de $\operatorname{Ker}(T)$, en este caso el polinomio constante 1; esto es,

$$\mathscr{B}_1 = \left\{x, \tfrac{1}{2}x^2, \tfrac{1}{3}x^3, 1\right\}.$$

Puesto que

$$\begin{aligned} T(x) &= 1 &&= 1\cdot 1 + 0\cdot x + 0\cdot x^2, \\ T\left(\tfrac{1}{2}x^2\right) &= x &&= 0\cdot 1 + 1\cdot x + 0\cdot x^2, \\ T\left(\tfrac{1}{3}x^3\right) &= x^2 &&= 0\cdot 1 + 0\cdot x + 1\cdot x^2, \\ T(1) &= 0 &&= 0\cdot 1 + 0\cdot x + 0\cdot x^2, \end{aligned}$$

se tiene

$$[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \blacktriangleleft$$

5.2.4 Isomorfismos

Entre dos idiomas diferentes no existe un diccionario perfecto con el cual se pueda traducir literalmente del uno al otro. La estructura lingüística es demasiado compleja, convencional, evolutiva y varía de un país a otro incluso entre dos regiones de un mismo país. Sin embargo, en las matemáticas es usual que existan "traductores" perfectos entre entes distintos cuando éstos tienen una estructura algebraica con un origen común. En álgebra lineal, cierto tipo de transformaciones lineales son los "traductores" que se emplean de manera natural entre distintos espacios vectoriales porque son funciones que preservan la estructura de las operaciones. En este breve apartado veremos cómo es posible construir estos traductores y probaremos que cualquier espacio vectorial de dimensión finita, salvo los sustantivos empleados para designar a sus elementos, es en esencia el espacio de vectores $\mathbb{R}^n$.

Definición 5.16 *Sean* $\mathbf{E}$ *y* $\mathbf{F}$ *un par de espacios vectoriales y* $T : \mathbf{E} \to \mathbf{F}$ *una transformación lineal.*

1. *T es un* ***isomorfismo***[18] *si T es una transformación biyectiva.*
2. *Si existe un isomorfismo del espacio* $\mathbf{E}$ *sobre el espacio* $\mathbf{F}$*, se dice que* $\mathbf{E}$ *es isomorfo a* $\mathbf{F}$ *y se escribe* $\mathbf{E} \cong F$.

[18] Del griego *iso-morfo* (igual forma o estructura).

▶ **Ejemplo 5.39** Sean $\mathbf{E} = \mathbb{R}^3$ y $\mathbf{F} = \mathbf{P}_2$, el espacio de polinomios de grado a lo más 2, y $T : \mathbf{E} \to \mathbf{F}$ definida por

$$T(a,b,c) = a + bx + cx^2.$$

Mostrar que T es un isomorfismo.◀

DEMOSTRACIÓN ■ Sean $(a,b,c), (a_1,b_1,c_1) \in \mathbb{R}^3$ y $\alpha, \beta \in \mathbb{R}$, entonces

$$\begin{aligned} T(\alpha(a,b,c) + \beta(a_1,b_1,c_1)) &= T(\alpha a + \beta a_1, \alpha b + \beta b_1, \alpha c + \beta c_1) \\ &= (\alpha a + \beta a_1) + (\alpha b + \beta b_1)x + (\alpha c + \beta c_1)x^2 \\ &= \alpha(a + bx + cx^2) + \beta(a_1 + b_1 x + c_1 x^2) \\ &= \alpha T(a,b,c) + \beta T(a_1,b_1,c_1); \end{aligned}$$

así que T es lineal. Si $(a,b,c) \in \text{Ker}(T)$, entonces $T(a,b,c) = 0$; y por tanto,

$$a + bx + cx^2 = 0$$

de donde $a = b = c = 0$, y por ende $\text{Ker}(T) = \{\vec{0}_{\mathbb{R}^3}\}$; luego T es inyectiva. Si $p(x) = a + bx + cx^2 \in \mathbf{F}$, claramente $T(a,b,c) = p(x)$; por lo que T es suprayectiva. Hemos mostrado que T es lineal y biyectiva, así que T es un isomorfismo. ■

A continuación probamos que $\cong$ es una relación de equivalencia entre los espacios vectoriales reales.

Teorema 5.21 *Sean* $\mathbf{E}, \mathbf{F}$ *y* $\mathbf{G}$ *espacios vectoriales. Entonces*

1. $\mathbf{E} \cong \mathbf{E}$.

2. $\mathbf{E} \cong \mathbf{F} \Rightarrow \mathbf{F} \cong \mathbf{E}$.

3. $\mathbf{E} \cong \mathbf{F}$, $\mathbf{F} \cong \mathbf{G} \Rightarrow \mathbf{E} \cong \mathbf{G}$.

DEMOSTRACIÓN ■

1. Sea $T : \mathbf{E} \to \mathbf{E}$ la transformación identidad; es decir, $T(\vec{u}) = \vec{u}\ \forall \vec{u} \in \mathbf{E}$. Claramente T es lineal y biyectiva.
2. Sea $T : \mathbf{E} \to \mathbf{F}$ un isomorfismo. Como T es biyectiva, entonces es invertible, y por el teorema 5.11, $T^{-1} : \mathbf{F} \to \mathbf{E}$ es lineal y biyectiva; luego $\mathbf{F} \cong \mathbf{E}$.
3. Se deja de ejercicio al lector. ■

❍ **Nota 5.8** Por el inciso 2 del teorema precedente, $\mathbf{E} \cong \mathbf{F} \Rightarrow \mathbf{F} \cong \mathbf{E}$, diremos simplemente que $\mathbf{E}$ y $\mathbf{F}$ son isomorfos si se cumple la propiedad 2 de la definición 5.16.

▶ **Ejemplo 5.40** Sean los espacios vectoriales $C[0,1]$ y $C[2,3]$. Mostrar que son isomorfos.◀

DEMOSTRACIÓN ■ Sea $T : C[0,1] \to C[2,3]$ la transformación definida por $T(f) = g$, donde $g(x) = f(x-2)\ \forall x \in [2,3]$. Geométricamente la transformación T asigna a cada función $f \in C[0,1]$ su traslación al intervalo $[2,3]$ como se ilustra en la figura 5-4.

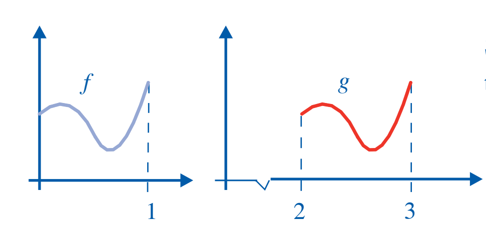


Figura 5-4 •

Sean $f_1, f_2 \in C[0,1]$ y $\alpha, \beta \in \mathbb{R}$. Entonces $\forall x \in [2,3]$ se tiene

$$\begin{aligned} T(\alpha f_1 + \beta f_2)(x) &= (\alpha f_1 + \beta f_2)(x-2) \\ &= \alpha f_1(x-2) + \beta f_2(x-2) \\ &= (\alpha T(f_1))(x) + (\beta T(f_2))(x); \end{aligned}$$

luego $T(\alpha f_1 + \beta f_2) = \alpha T(f_1) + \beta T(f_2)$ lo cual prueba que T es lineal. Sea $f \in \operatorname{Ker}(T)$, entonces $T(f) = \theta$, el neutro aditivo de $C[2,3]$, por lo que

$$f(x-2) = 0 \quad \forall x \in [2,3].$$

Sea $u \in [0,1]$, entonces $x = u + 2 \in [2,3]$, así que

$$f(u) = f(x-2) = 0,$$

y por tanto, $f(u) = 0$ para todo $u \in [0,1]$; i.e., $f = \theta$, el neutro aditivo de $C[0,1]$; luego $\operatorname{Ker}(T) = \{\theta\}$ y entonces T es inyectiva. Sean $g \in C[2,3]$ y $f : [0,1] \to \mathbb{R}$ la función definida por $f(x) = g(x+2)$ (puesto que $0 \le x \le 1$, se tiene $2 \le x+2 \le 3$). Entonces, ya que g y la función $x \mapsto x+2$ son continuas, f es continua en $[0,1]$; y para todo $u \in [2,3]$

$$T(f)(u) = f(u-2) = g(u)$$

es decir,

$$T(f) = g;$$

lo cual prueba que T es suprayectiva. Con esto hemos demostrado que T es un isomorfismo y por tanto $C[0,1] \cong C[2,3]$. ■

Del inciso 5 del teorema 5.10 (cfr. pág. 432) se tiene como consecuencia inmediata el siguiente resultado.

Teorema 5.22 *Si* **E** *y* **F** *son espacios vectoriales isomorfos y uno de ellos tiene dimensión finita n, entonces el otro tiene también dimensión n.*

▶ **Ejemplo 5.41** Sean **E** y **F** un par de espacios vectoriales reales de dimensiones n y m, respectivamente. Entonces $\mathscr{L}(\mathbf{E},\mathbf{F}) \cong \mathfrak{M}_{m\times n}$. En efecto, sean $\mathscr{B}_1 = \{\vec{e}_1, \ldots, \vec{e}_n\}$ y $\mathscr{B}_2 = \{\vec{f}_1, \ldots, \vec{f}_m\}$ bases de **E** y **F**, respectivamente. Para cada $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$ sea $[T]_{\mathscr{B}_1}^{\mathscr{B}_2}$ la representación matricial de T relativa a estas bases. Definimos $\Psi : \mathscr{L}(\mathbf{E},\mathbf{F}) \to \mathfrak{M}_{m\times n}$ como $\Psi(T) = [T]_{\mathscr{B}_1}^{\mathscr{B}_2}$. Entonces:

1. Sean $\alpha, \beta \in \mathbb{R}$, $T_1, T_2 \in \mathscr{L}(\mathbf{E},\mathbf{F})$. Para todo $\vec{u} \in \mathbf{E}$ se tiene, por 5.12 y 5.12 del teorema 5.12, pág. 436,

$$\begin{aligned} (\alpha[T_1]_{\mathscr{B}_1}^{\mathscr{B}_2} + \beta[T_2]_{\mathscr{B}_1}^{\mathscr{B}_2})[\vec{u}]_{\mathscr{B}_1} &= \alpha[T_1]_{\mathscr{B}_1}^{\mathscr{B}_2}[\vec{u}]_{\mathscr{B}_1} + \beta[T_2]_{\mathscr{B}_1}^{\mathscr{B}_2}[\vec{u}]_{\mathscr{B}_1} \\ &= \alpha[T_1(\vec{u})]_{\mathscr{B}_2} + \beta[T_2(\vec{u})]_{\mathscr{B}_2} \\ &= [\alpha T_1(\vec{u}) + \beta T_2(\vec{u})]_{\mathscr{B}_2}. \end{aligned}$$

De la unicidad de la representación matricial relativa a bases dadas de una transformación lineal (cfr. nota 5.7, pág. 448), se tiene

$$[\alpha T_1+\beta T_2]_{\mathscr{B}_1}^{\mathscr{B}_2}=\alpha[T_1]_{\mathscr{B}_1}^{\mathscr{B}_2}+\beta[T_2]_{\mathscr{B}_1}^{\mathscr{B}_2};$$

es decir,

$$\Psi(\alpha T_1+\beta T_2)=\alpha\Psi(T_1)+\beta\Psi(T_2).$$

2. Sea $T\in\operatorname{Ker}(\Psi)$, entonces $\Psi(T)=\mathscr{O}$, la matriz cero de tamaño $m\times n$. Es decir,

$$[T]_{\mathscr{B}_1}^{\mathscr{B}_2}=\mathscr{O};$$

por tanto, para todo $\vec{u}\in\mathbf{E}$,

$$[T]_{\mathscr{B}_1}^{\mathscr{B}_2}[\vec{u}]_{\mathscr{B}_1}=\vec{0}_{\mathbb{R}^m},$$

luego

$$[T(\vec{u})]_{\mathscr{B}_2}=\vec{0}_{\mathbb{R}^m}.$$

De 4 del teorema 5.12, la igualdad precedente implica

$$T(\vec{u})=\vec{0}_{\mathbf{F}}\quad\forall\vec{u}\in\mathbf{E}.$$

Es decir, T es la transformación cero, el neutro aditivo de $\mathscr{L}(\mathbf{E},\mathbf{F})$; lo cual prueba que Ψ es inyectiva.

3. Sea $A\in\mathfrak{M}_{m\times n}$. Definimos $T:\mathbf{E}\to\mathbf{F}$, para cada $\vec{u}\in\mathbf{E}$, por

$$T(\vec{u})=\alpha_1\vec{f}_1+\cdots+\alpha_m\vec{f}_m,$$

donde $(\alpha_1,\ldots,\alpha_m)=A[\vec{u}]_{\mathscr{B}_1}$. Dado que $\vec{u}=\vec{v}\Leftrightarrow[\vec{u}]_{\mathscr{B}_1}=[\vec{v}]_{\mathscr{B}_1}$, se desprende que T está bien definida. Sean $\alpha,\beta\in\mathbb{R}$ y $\vec{u},\vec{v}\in\mathbf{E}$, entonces

$$[\alpha\vec{u}+\beta\vec{v}]_{\mathscr{B}_1}=\alpha[\vec{u}]_{\mathscr{B}_1}+\beta[\vec{v}]_{\mathscr{B}_1},$$

y por tanto

$$A[\alpha\vec{u}+\beta\vec{v}]_{\mathscr{B}_1}=\alpha A[\vec{u}]_{\mathscr{B}_1}+\beta A[\vec{v}]_{\mathscr{B}_1};$$

de donde

$$T(\alpha\vec{u}+\beta\vec{v})=\alpha T(\vec{u})+\beta T(\vec{v}).$$

Luego $T\in\mathscr{L}(\mathbf{E},\mathbf{F})$. Puesto que

$$[T(\vec{u})]_{\mathscr{B}_2}=A[\vec{u}]_{\mathscr{B}_1}\quad\forall\vec{u}\in\mathbf{E},$$

se tiene $A=[T]_{\mathscr{B}_1}^{\mathscr{B}_2}$; esto es

$$\Psi(T) = A.$$

Lo cual prueba que T es suprayectiva.

Hemos mostrado así que Ψ es un isomorfismo y por tanto $\mathscr{L}(\mathbf{E},\mathbf{F}) \cong \mathfrak{M}_{m\times n}$.◀

Como consecuencia del ejemplo anterior y del teorema 5.22,

$$\dim(\mathscr{L}(\mathbf{E},\mathbf{F})) = mn$$

si $\dim(\mathbf{E}) = n$ y $\dim(\mathbf{F}) = m$.

Ahora mostraremos que cualquier espacio vectorial real de dimensión finita es isomorfo a $\mathbb{R}^n$.

Teorema 5.23 *Sea* $\mathbf{E}$ *un espacio vectorial real de dimensión n, entonces* $\mathbf{E} \cong \mathbb{R}^n$.

DEMOSTRACIÓN ■ Sea $\mathscr{B} = \{\vec{e}_1,\ldots,\vec{e}_n\}$ una base de $\mathbf{E}$. Sea $T : \mathbf{E} \to \mathbb{R}^n$ definida como

$$T(\vec{u}) = [\vec{u}]_{\mathscr{B}}.$$

Entonces, por las propiedades 1 y 2 del teorema 5.12 se tiene

$$\begin{aligned} T(\alpha\vec{u}+\beta\vec{v}) &= [\alpha\vec{u}+\beta\vec{v}]_{\mathscr{B}} \\ &= \alpha[\vec{u}]_{\mathscr{B}}+\beta[\vec{v}]_{\mathscr{B}} \\ &= \alpha T(\vec{u})+\beta T(\vec{v}); \end{aligned}$$

por lo que T es lineal. Si $(\alpha_1,\ldots,\alpha_n) \in \mathbb{R}^n$, sea $\vec{u} = \alpha_1\vec{e}_1+\cdots+\alpha_n\vec{e}_n \in \mathbf{E}$, entonces

$$T(\alpha_1\vec{e}_1+\cdots+\alpha_n\vec{e}_n) = (\alpha_1,\ldots,\alpha_n);$$

así que T es suprayectiva. Si $T(\vec{u}) = \vec{0}_{\mathbb{R}^n}$, entonces $[\vec{u}]_{\mathscr{B}} = \vec{0}_{\mathbb{R}^n}$ y por la propiedad 5.12 del teorema 5.12 se desprende $\vec{u} = \vec{0}_{\mathbf{E}}$; lo cual demuestra que T es inyectiva. Hemos probado así que T es un isomorfismo, y por tanto $\mathbf{E} \cong \mathbb{R}^n$. ■

Si $\mathbf{E}$ y $\mathbf{F}$ son espacios vectoriales reales con dimensiones iguales a n, entonces, por el teorema precedente, $\mathbf{E} \cong \mathbb{R}^n$ y $\mathbf{F} \cong \mathbb{R}^n$; por tanto, del tercer inciso del teorema 5.21, $\mathbf{E} \cong \mathbf{F}$. Con lo que queda demostrado el siguiente teorema.

Teorema 5.24 *Cualquier par de espacios vectoriales que tienen la misma dimensión (finita) son isomorfos.*

▶ Ejemplo 5.42

1. $\mathbf{P}_n \cong \mathbb{R}^{n+1}$.
2. $\mathfrak{M}_{m\times n} \cong \mathbb{R}^{mn}$.
3. $\mathscr{L}(\mathbf{P_{n-1}},\mathbf{P_{n-1}}) \cong \mathfrak{M}_n$ (las matrices cuadradas de orden n).◀

5.3 Valores y vectores propios, diagonalización

Las propiedades de un operador lineal T, definido en un espacio de dimensión finita, que son independientes del sistema de coordenadas (la base) con el que se trabaje en el espacio (es decir, que son independientes de la representación matricial), se denominan *propiedades intrínsecas* del operador; por ejemplo, el determinante de cualquier representación matricial de un operador lineal es independiente de la representación y, por tanto, es una propiedad intrínseca de éste. Entonces, como mencionamos antes, es sumamente importante encontrar alguna representación matricial lo suficientemente sencilla para identificar, a partir de esta representación, propiedades que sean intrínsecas del operador. En la sección anterior probamos, en el teorema 5.20, que dada una transformación lineal $T : \mathbf{E} \to \mathbf{F}$, se pueden construir bases $\{\vec{e}_i\}$ y $\{\vec{f}_j\}$ de tal suerte que la representación matricial de T relativa a estas bases tiene unos en la diagonal principal y ceros fuera de ella; es decir, una representación *diagonal*[19] en ese sentido. Para operadores lineales las representaciones matriciales se definen a partir de una sola base $\{\vec{e}_i\}$ para el espacio vectorial; en este caso no siempre es posible construir una representación matricial diagonal sin emplear otra base más. El objetivo ahora, en esta sección, es determinar qué tipo de operadores tienen representaciones matriciales (relativas a una sola base para el espacio) que sean diagonales y cómo encontrar dichas representaciones.

5.3.1 Valores y vectores propios

Supongamos que $\mathbf{E}$ es un espacio con dimensión finita n y $T : \mathbf{E} \to \mathbf{E}$ es un operador lineal que tiene una representación matricial $[T]_{\mathscr{B}}$, relativa a la base $\mathscr{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$, que es una matriz diagonal. Sean $a_{ii} = \lambda_i$, $i = 1, \ldots, n$, los elementos de la diagonal de la matriz $[T]_{\mathscr{B}}$. Entonces, dado que

$$[\vec{e}_i]_{\mathscr{B}} = (0, \ldots, 0, \overset{i}{1}, 0, \ldots, 0),$$

para $i = 1, \ldots, n$, y[20] $[T]_{\mathscr{B}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$,

$$\begin{aligned} [T(\vec{e}_i)]_{\mathscr{B}} &= [T]_{\mathscr{B}} [\vec{e}_i]_{\mathscr{B}} \\ &= \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \lambda_i \\ 0 \\ \vdots \\ 0 \end{bmatrix} \ (\to \text{fila } i\ ). \end{aligned}$$

Por tanto,

$$T(\vec{e}_i) = \lambda_i \vec{e}_i \quad \text{para cada } i = 1, \ldots, n. \tag{5.19}$$

[19] Recordemos que una matriz *cuadrada* es diagonal si todas las componentes fuera de la diagonal son nulas. Abusando un poco del lenguaje, entenderemos que una matriz *no cuadrada* es diagonal si todas las componentes fuera de la diagonal principal son nulas; donde la diagonal principal está formada por las entradas de la forma a_{ii}.

[20] Cfr. definición 1.4, página 8.

Ahora supongamos que existe una base $\mathscr{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ del espacio $\mathbf{E}$ y escalares $\lambda_1, \ldots, \lambda_n$ tales que se cumple la igualdad (5.19). Entonces

$$T(\vec{e}_i) = \lambda_i \vec{e}_i = 0 \cdot \vec{e}_1 + \cdots + 0 \cdot \vec{e}_{i-1} + \lambda_i \cdot \vec{e}_i + 0 \cdot \vec{e}_{i+1} + \cdots + 0 \cdot \vec{e}_n$$

para cada $i = 1, \ldots, n$; luego

$$[T]_{\mathscr{B}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Esto es, para que un operador lineal T tenga una representación matricial diagonal es necesario y suficiente que existan n-vectores L.I., $\vec{e}_1, \ldots, \vec{e}_n$, y n-escalares $\lambda_1, \ldots, \lambda_n$ que cumplan la igualdad (5.19). Resumimos lo hasta aquí demostrado en el siguiente teorema.

Teorema 5.25 *Sean $\mathbf{E}$ un espacio vectorial de dimensión n y $T : \mathbf{E} \to \mathbf{E}$ un operador lineal. Entonces T tiene una representación matricial diagonal si y sólo si existen $\vec{e}_1, \ldots, \vec{e}_n \in \mathbf{E}$, vectores L.I., y un correspondiente conjunto de escalares $\lambda_1, \ldots, \lambda_n$, tales que*

$$T(\vec{e}_i) = \lambda_i \vec{e}_i \quad \textit{para cada } i = 1, \ldots, n.$$

En ese caso la representación diagonal es

$$[T]_{\mathscr{B}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

donde $\mathscr{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$.

De esta manera, el problema de hallar representaciones matriciales diagonales de operadores lineales se reduce a encontrar los vectores L.I. y los correspondientes escalares que satisfacen (5.19). El primer paso para resolver este tipo de problemas es encontrar escalares λ y vectores correspondientes $\vec{u}$ tales que $T(\vec{u}) = \lambda \vec{u}$. Motivados con este primer objetivo damos la siguiente definición.

Definición 5.17 ***(Valores propios de operadores lineales)*** *Sean $\mathbf{E}$ un espacio vectorial (no necesariamente de dimensión finita) y $T : \mathbf{E} \to \mathbf{E}$ un operador lineal.*

1. *Se dice que $\lambda \in \mathbb{R}$ es un **valor propio*** (valor característico, eigenvalor, autovalor) *de T si existe un vector $\vec{u} \in \mathbf{E}$, con $\vec{u} \neq \vec{0}_{\mathbf{E}}$, tal que*

$$T(\vec{u}) = \lambda \vec{u} \tag{5.20}$$

2. *Si λ es un valor propio de T y $\vec{u} \in \mathbf{E}$ es un vector no nulo que satisface (5.20), entonces se dice que $\vec{u}$ es un **vector propio*** (vector característico, eigenvector, autovector) *del operador T correspondiente al valor propio λ.*

❍ Nota 5.9

1. Si λ es valor propio del operador lineal T, entonces existe $\vec{u} \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$ que satisface (5.20); luego, si $c \in \mathbb{R} - \{0\}$ es un escalar, $T(c\vec{u}) = cT(\vec{u}) = c(\lambda \vec{u}) = \lambda(c\vec{u})$ y, por tanto, $c\vec{u} \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$ es vector propio de T correspondiente a λ. Así que un valor propio tiene una infinidad de vectores propios correspondientes.

2. Supongamos que λ, μ son valores propios con un mismo vector propio correspondiente $\vec{u}$, entonces $\lambda\vec{u} = T(\vec{u}) = \mu\vec{u}$, lo cual implica $(\lambda - \mu)\vec{u} = \vec{0}_{\mathbf{E}}$, y por ende $\lambda = \mu$ (pues $\vec{u} \neq \vec{0}_{\mathbf{E}}$, cfr. la propiedad 9 del teorema 3.4, pág. 138). Es decir, existe un único valor propio correspondiente a un vector propio dado.
3. Si en la definición 5.17 se permitiera que $\vec{u} = \vec{0}_{\mathbf{E}}$, entonces todo escalar λ sería valor propio de T; pues $\vec{0}_{\mathbf{E}} = T(\vec{0}_{\mathbf{E}}) = \lambda\vec{0}_{\mathbf{E}}$. El restringir los vectores propios a vectores no nulos es porque se quieren encontrar vectores L.I. que satisfagan (5.20), y si $\vec{0}_{\mathbf{E}}$ se incluye en esta lista entonces este conjunto se convierte en linealmente dependiente. Otra razón es que se rompe la unicidad de valores propios correspondientes a un vector propio dado que se mostró en el segundo inciso.
4. Aunque los vectores propios no pueden ser nulos, los valores propios sí. Por ejemplo, si T es un operador lineal no inyectivo, entonces $T(\vec{u}) = \vec{0}_{\mathbf{E}} = 0 \cdot \vec{u}$ para todo $\vec{u} \in \operatorname{Ker}(T)$; puesto que T no es inyectivo, el núcleo contiene elementos no nulos y así $\lambda = 0$ es valor propio de T.
5. Es evidente que λ es valor propio del operador T si y sólo si $\operatorname{Ker}(T - \lambda I) \neq \{\vec{0}_{\mathbf{E}}\}$; es decir, si el operador lineal $T - \lambda I$ es no inyectivo, donde I es el operador identidad del espacio **E** (cfr. ejemplo 5.8, pág. 418).

▶ **Ejemplo 5.43** Sea el espacio vectorial

$$C^{\infty}(-\infty,\infty) = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ tiene derivada de todo orden para cada } x \in \mathbb{R}\}$$

y sea el operador lineal $T : C^{\infty}(-\infty,\infty) \to C^{\infty}(-\infty,\infty)$, $T(f) = f'$, el operador derivación. Encontrar los valores propios y vectores propios correspondientes de T.◀

Solución 1. Si $\lambda = 0$

$$\begin{aligned} T(f) &= \lambda f &= 0 \cdot f = \theta \Rightarrow \\ f' &= \theta &\Rightarrow \\ f'(x) &= 0 &\forall x; \end{aligned}$$

por tanto, f es una constante. Las funciones constantes distintas de cero son los vectores propios correspondientes al valor propio $\lambda = 0$.

2. Si $\lambda \neq 0$

$$\begin{aligned} T(f) &= \lambda f \Rightarrow \\ f' &= \lambda f \Rightarrow \\ \frac{df}{dx} &= \lambda f \Rightarrow \\ \int \frac{df}{f} &= \lambda \int dx \Rightarrow \\ \ln|f| &= \lambda x + c \Rightarrow \\ f &= ke^{\lambda x} \qquad (k = \pm e^{c}). \end{aligned}$$

Vectores propios para λ:

$$f(x) = ke^{\lambda x};\ k \neq 0.$$

En síntesis: todo número real λ es valor propio de T. Si $\lambda = 0$ los vectores propios correspondientes son todas las funciones constantes no nulas. Si $\lambda \neq 0$, los vectores propios correspondientes a λ son las funciones $f(x) = ke^{\lambda x}$, $k \neq 0$. ✔

No todo operador lineal tiene valores propios, como hacemos patente en el siguiente ejemplo.

► **Ejemplo 5.44** (Un operador sin valores propios) Sea $T : C[0,1] \to C[0,1]$ el operador definido por $T(f) = g$, donde

$$g(x) = \int_0^x f(t)dt, \quad x \in [a,b].$$

Por ejemplo, si $f(x) = x^2$, entonces $g(x) = \int_0^x t^2 dt = \frac{x^3}{3}$. Encontrar, si es que existen, los valores propios de T.◄

Solución Supongamos que λ es un valor propio de T. Si $f \neq \theta$ es una función propia de λ se debe tener

$$T(f) = \lambda f.$$

Luego,

$$\lambda f(x) = \int_0^x f(t)dt \quad \forall x \in [0,1]. \tag{5.21}$$

Entonces,

$$(\lambda f(x))' = \left(\int_0^x f(t)dt\right)'$$

y al aplicar el teorema fundamental del cálculo[21] se obtiene

$$f(x) = \lambda f'(x).$$

De aquí que λ debe ser distinto de cero, pues en caso contrario $f = \theta$ (la función constante cero). La igualdad precedente se puede escribir como

$$\frac{df}{dx} = \frac{1}{\lambda} f$$

que equivale a

$$\frac{df}{f} = \frac{1}{\lambda} dx$$

e integrando se tiene

$$\ln|f| = \frac{1}{\lambda} x + C,$$

[21] Si $f \in C[a,b]$ y $F(x) = \int_a^x f(t)dt$, entonces F es derivable y $F'(x) = f(x)\ \forall x \in [a,b]$.

que al tomar exponencial en ambos lados produce

$$f(x) = ke^{x/\lambda},$$

con $k = \pm e^C \neq 0$. Sin embargo, la función $f(x) = ke^{x/\lambda}$ no satisface el requerimiento (5.21):

$$\lambda k = \lambda f(0) = \int_0^0 f(t)dt = 0$$

pues $\lambda k \neq 0$. Luego T no tiene valores propios. ✔

Espacio propio

Sea $T : \mathbf{E} \to \mathbf{E}$ un operador lineal y sea λ un valor propio de T; si $\vec{u}, \vec{v} \in \mathbf{E}$ son vectores propios correspondientes a λ y $\alpha, \beta \in \mathbb{R}$, entonces

$$\begin{aligned} T(\alpha\vec{u} + \beta\vec{v}) &= \alpha T(\vec{u}) + \beta T(\vec{v}) \\ &= \alpha(\lambda\vec{u}) + \beta(\lambda\vec{v}) \\ &= \lambda(\alpha\vec{u} + \beta\vec{v}). \end{aligned}$$

Por lo que si $\alpha\vec{u} + \beta\vec{v} \neq \vec{0}_{\mathbf{E}}$, éste es también un vector propio correspondiente al valor propio λ. Así, el conjunto de todos los vectores $\vec{u}$ tales que $T(\vec{u}) = \lambda\vec{u}$ es un subespacio vectorial de $\mathbf{E}$ (este conjunto también contiene a $\vec{0}_{\mathbf{E}}$, ya que $T(\vec{0})_{\mathbf{E}} = \vec{0}_{\mathbf{E}} = \lambda\vec{0}_{\mathbf{E}}$) que contiene a todos los vectores propios correspondientes a λ. Hacemos patente este concepto en la siguiente definición.

Definición 5.18 *Sean $\mathbf{E}$ un espacio vectorial, $T : \mathbf{E} \to \mathbf{E}$ un operador lineal y $\lambda \in \mathbb{R}$ un valor propio de T. Al subespacio vectorial de $\mathbf{E}$*

$$E_\lambda = \{\vec{u} \in \mathbf{E} \,|\, T(\vec{u}) = \lambda\vec{u}\}$$

*se le llama el **espacio propio** del valor propio λ.*

Es decir, el espacio propio de un valor propio λ es el conjunto de todos los vectores propios correspondientes a λ uniendo al mismo el vector neutro aditivo del espacio.

Sea A una matriz cuadrada de orden n y consideremos el operador matricial $T_A : \mathbb{R}^n \to \mathbb{R}^n$, $T_A(\vec{u}) = A\vec{u}$. Entonces λ es valor propio de T_A si y sólo si existe $\vec{u} \in \mathbb{R}^n - \vec{0}_{\mathbb{R}^n}$ tal que $A\vec{u} = \lambda\vec{u}$. De esta manera, la definición 5.17 tiene como caso particular la siguiente definición.

Definición 5.19 ***(Valores propios de matrices)*** *Sea A una matriz cuadrada de orden n. Se dice que $\lambda \in \mathbb{R}$ es un **valor propio** (valor característico, autovalor o eigenvalor) de A si existe $\vec{u} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$ tal que*

$$A\vec{u} = \lambda\vec{u}.$$

*En tal caso a $\vec{u}$ se le llama **vector propio** (vector característico, autovector o eigenvector) de la matriz A correspondiente a λ.*

▶ **Ejemplo 5.45** Si $A = \begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}$,

$$\begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Así, 4 es valor propio de A.◀

▶ **Ejemplo 5.46** Sea A la matriz del ejemplo precedente. Encontrar los vectores propios correspondientes al valor propio 4 y E_4 (el espacio propio del valor propio 4).◀

Solución Los vectores propios de A, correspondientes a este valor propio, son las soluciones no triviales del sistema

$$A\vec{u} = 4\vec{u}$$

que equivalen a las soluciones no triviales del sistema homogéneo

$$(A - 4I_2)\,\vec{u} = \vec{0}_{\mathbb{R}^2}.$$

Resolvamos entonces este sistema:

$$\begin{bmatrix} 2-4 & 2 \\ 3 & 1-4 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}.$$

Luego

$$\vec{u} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r \\ r \end{bmatrix} = r\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad r \neq 0.$$

Es decir, el conjunto de vectores propios correspondientes al valor propio 4 son todos los vectores de la forma $r(1,1)$ con $r \neq 0$ y el espacio propio es $E_4 = \text{gn}((1,1))$. ✔

Teorema 5.26 *Sean A, B matrices cuadradas de orden n que son similares. Entonces A y B tienen los mismos valores propios.*

DEMOSTRACIÓN ■ Sean λ un valor propio de la matriz A, $\vec{u} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$ un vector propio correspondiente, y P una matriz cuadrada de orden n tal que $A = P^{-1}BP$. Entonces

$$B(P\vec{u}) = P(A\vec{u}) = P(\lambda\vec{u}) = \lambda\,(P\vec{u}).$$

Puesto que $\vec{u} \neq \vec{0}_{\mathbb{R}^n}$ y P es una matriz invertible, se tiene que $P\vec{u} \neq \vec{0}_{\mathbb{R}^n}$; por tanto, λ es valor propio de B. Sea μ un valor propio de B y $\vec{v} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$ un vector propio correspondiente. Entonces

$$A(P^{-1}\vec{v}) = P^{-1}(B\vec{v}) = P^{-1}(\mu\vec{v}) = \mu(P^{-1}\vec{v})$$

y nuevamente, dado que P^{-1} es invertible y $\vec{v} \neq \vec{0}_{\mathbb{R}^n}$, $P^{-1}\vec{v} \neq \vec{0}_{\mathbb{R}^n}$ y, por tanto μ es valor propio de B. ■

Ya que en un espacio vectorial de dimensión finita cualquier par de representaciones matriciales de un operador lineal son similares, el teorema precedente tiene como consecuencia inmediata el siguiente corolario.

Corolario 5.1 *Si* $\mathbf{E}$ *es un espacio de dimensión finita y* $T:\mathbf{E}\to\mathbf{E}$ *es un operador lineal, entonces todas las representaciones matriciales de* T *tienen los mismos valores propios.*

¿Cuál es entonces la relación entre los valores propios de un operador lineal y los valores propios de cualquiera de sus representaciones matriciales? Seguramente el lector ya sabe cuál es la respuesta; pero para hacerla patente la probamos en la siguiente proposición.

Teorema 5.27 *Sean* $\mathbf{E}$ *un espacio vectorial de dimensión finita* n *y* $T:\mathbf{E}\to\mathbf{E}$ *un operador lineal. Si* $[T]_{\mathscr{B}}$ *es cualquier representación matricial de* T *relativa a una base* $\mathscr{B}=\{\vec{e}_1,\ldots,\vec{e}_n\}$*, entonces el operador* T *y la matriz* $[T]_{\mathscr{B}}$ *tienen los mismos valores propios.*

DEMOSTRACIÓN ■ Supongamos que λ es un valor propio de T con $\vec{u}\in\mathbf{E}-\{\vec{0}_{\mathbf{E}}\}$ un vector propio correspondiente, entonces $T(\vec{u})=\lambda\vec{u}$. Por tanto,

$$[\lambda\vec{u}]_{\mathscr{B}}=[T(\vec{u})]_{\mathscr{B}}=[T]_{\mathscr{B}}[\vec{u}]_{\mathscr{B}}.$$

Luego

$$[T]_{\mathscr{B}}[\vec{u}]_{\mathscr{B}}=\lambda[\vec{u}]_{\mathscr{B}}$$

y dado que[22] $\vec{u}\neq\vec{0}_{\mathbf{E}}\Rightarrow[\vec{u}]_{\mathscr{B}}\neq\vec{0}_{\mathbb{R}^n}$, se tiene que λ es valor propio de la matriz $[T]_{\mathscr{B}}$. Ahora supongamos que μ es un valor propio de la matriz $[T]_{\mathscr{B}}$, con un vector propio correspondiente $\vec{v}=(v_1,\ldots,v_n)\in\mathbb{R}^n-\{\vec{0}_{\mathbb{R}^n}\}$. Sea $\vec{u}=v_1\vec{e}_1+\cdots+v_n\vec{e}_n$, entonces[23] $\vec{u}\in\mathbf{E}-\{\vec{0}_{\mathbf{E}}\}$ y

$$\begin{aligned}[T(\vec{u})]_{\mathscr{B}}&=[T]_{\mathscr{B}}[\vec{u}]_{\mathscr{B}}\\&=[T]_{\mathscr{B}}\vec{v}\\&=\mu\vec{v}\\&=\mu[\vec{u}]_{\mathscr{B}}\\&=[\mu\vec{u}]_{\mathscr{B}}\end{aligned}$$

lo cual implica

$$T(\vec{u})=\mu\vec{u}$$

y por ende μ es valor propio de T. ■

[22] Cfr. teorema 5.12, página 436.
[23] Cfr. teorema 5.12, página 436.

En resumen, si $T : \mathbf{E} \to \mathbf{E}$ es un operador lineal y $\mathbf{E}$ tiene dimensión finita, los valores propios de T son los valores propios de cualquier representación matricial del mismo. Por tanto, el problema de hallar los valores propios de un operador lineal definido en un espacio de dimensión finita se reduce a encontrar los valores propios de una matriz (de una representación matricial cualquiera del mismo). Esta es una enorme ventaja, pues podemos aprovechar toda la herramienta matricial que hemos desarrollado a lo largo de este libro para este fin. Por esta razón es que será suficiente restringir nuestro estudio de encontrar valores propios al caso de matrices cuadradas de orden n.

Cálculo de valores y vectores propios de matrices

En el ejemplo 5.46 vimos cómo calcular los valores propios de una matriz cuadrada resolviendo un sistema homogéneo. Ésta es la idea general para calcular valores y vectores propios de una matriz cuadrada. Sea A una matriz cuadrada de orden n; dado que un sistema cuadrado homogéneo tiene soluciones no triviales si y sólo si la matriz de coeficientes del mismo tiene determinante cero, tenemos:

$$\begin{aligned}\lambda \text{ es valor propio de } A &\Leftrightarrow \text{existe } \vec{u} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\} \text{ tal que } A\vec{u} = \lambda\vec{u} \\ &\Leftrightarrow \text{el sistema homogéneo } (A - \lambda I_n)\vec{x} = \vec{0}_{\mathbb{R}^n} \\ &\quad\ \text{tiene soluciones no triviales} \\ &\Leftrightarrow \det(A - \lambda I_n) = 0\end{aligned}$$

donde I_n es la matriz identidad de orden n. Ahora, si $\lambda \in \mathbb{R}$ es un valor propio de la matriz A, entonces debe existir una solución no trivial del sistema homogéneo

$$(A - \lambda I_n)\vec{x} = \vec{0}_{\mathbb{R}^n}.$$

Así, los vectores propios correspondientes a λ son las soluciones no triviales de este sistema homogéneo. Hemos probado así el siguiente resultado.

Teorema 5.28 *Sea A una matriz cuadrada de orden n. Entonces*

1. *Las soluciones reales de la ecuación*

$$\det(A - \lambda I_n) = 0$$

son los valores propios de A.

2. *Y si $\lambda \in \mathbb{R}$ es un valor propio de A, las soluciones no triviales del sistema homogéneo*

$$(A - \lambda I_n)\vec{x} = \vec{0}_{\mathbb{R}^n} \tag{5.22}$$

son los vectores propios correspondientes a λ.

► **Ejemplo 5.47** Sea

$$A = \begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}.$$

Encontrar sus valores propios y los vectores propios correspondientes.◄

Solución

$$\begin{aligned}\det(A-\lambda I_2) &= \begin{vmatrix} 2-\lambda & 2 \\ 3 & 1-\lambda \end{vmatrix} \\ &= (2-\lambda)(1-\lambda)-6 \\ &= \lambda^2-3\lambda-4 \\ &= (\lambda-4)(\lambda+1)\end{aligned}$$

y las soluciones de

$$\det(A-\lambda I_2)=0$$

son los valores propios de A:

$$\lambda=4 \text{ y } \lambda=-1.$$

Si $\lambda=4$, el sistema homogéneo (5.22) en este caso es

$$\begin{bmatrix} 2-4 & 2 \\ 3 & 1-4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix};$$

que podemos resolver por el método de Gauss llevando a forma escalonada

$$\begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix}\sim\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

y haciendo sustitución regresiva

$$\begin{aligned}\begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} r \\ r \end{bmatrix} \\ &= r\begin{bmatrix} 1 \\ 1 \end{bmatrix}.\end{aligned}$$

Así, los vectores propios correspondientes a $\lambda=4$ son todos los vectores de la forma $r(1,1)$ con $r\neq 0$. Análogamente si $\lambda=-1$, el sistema homogéneo (5.22) para este caso es

$$\begin{bmatrix} 2+1 & 2 \\ 3 & 1+1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

y

$$\begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix}\sim\begin{bmatrix} 3 & 2 \\ 0 & 0 \end{bmatrix}$$

por tanto,

$$\begin{aligned}\begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} -(2/3)s \\ s \end{bmatrix} \\ &= \begin{bmatrix} -(2/3)3t \\ 3t \end{bmatrix} \\ &= t\begin{bmatrix} -2 \\ 3 \end{bmatrix}.\end{aligned}$$

Luego los vectores propios correspondientes a $\lambda=-1$ son todos los vectores de la forma $t(-2,3)$ con $t\neq 0$. ✔

En el ejemplo precedente obtuvimos $\det(A-\lambda I_2)=\lambda^2-3\lambda-4$ para la matriz $A=\begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}$. En el caso general, si se desarrolla $\det(A-\lambda I_n)$ para una matriz cuadrada A de orden n, se obtiene un polinomio en λ de grado n. Veamos algunos casos particulares:

1. Si $n=2$

$$\begin{aligned}\det(A-\lambda I_2) &= \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22-\lambda} \end{vmatrix} \\ &= (a_{11}-\lambda)(a_{22}-\lambda)-a_{12}a_{21} \\ &= \lambda^2-(a_{11}+a_{22})\lambda+a_{11}a_{22}-a_{12}a_{21}.\end{aligned}$$

2. Si $n=3$, empleando el inciso 1 tenemos

$$\begin{aligned}\det(A-\lambda I_3) &= \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix} \\ &= (a_{11}-\lambda)\begin{vmatrix} a_{22}-\lambda & a_{23} \\ a_{32} & a_{33}-\lambda \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33}-\lambda \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22}-\lambda & a_{23} \end{vmatrix} \\ &= (a_{11}-\lambda)\left(\lambda^2-(a_{22}+a_{33})\lambda+a_{22}a_{33}-a_{23}a_{32}\right) \\ &\quad -a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33}-\lambda \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22}-\lambda & a_{23} \end{vmatrix}\end{aligned}$$

y ya que los dos últimos términos son claramente polinomios de grado a lo más uno, se deduce que

$$\det(A-\lambda I_3) = -\lambda^3+(a_{22}+a_{33}+a_{11})\lambda^2+\alpha_1\lambda+\alpha_0$$

donde $\alpha_1, \alpha_0 \in \mathbb{R}$.

3. Si $n=4$, empleando los incisos anteriores,

$$\begin{aligned}\det(A-\lambda I_4) &= \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22}-\lambda & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33}-\lambda & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44}-\lambda \end{vmatrix} \\ &= (a_{11}-\lambda)\begin{vmatrix} a_{22}-\lambda & a_{23} & a_{24} \\ a_{32} & a_{33}-\lambda & a_{34} \\ a_{42} & a_{43} & a_{44}-\lambda \end{vmatrix} \\ &\quad -a_{21}\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33}-\lambda & a_{34} \\ a_{42} & a_{43} & a_{44}-\lambda \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22}-\lambda & a_{23} & a_{24} \\ a_{42} & a_{43} & a_{44}-\lambda \end{vmatrix} \\ &\quad -a_{41}\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22}-\lambda & a_{23} & a_{24} \\ a_{32} & a_{33}-\lambda & a_{34} \end{vmatrix} \\ &= (a_{11}-\lambda)(-\lambda^3+(a_{22}+a_{33}+a_{44})\lambda^2+\alpha_1\lambda+\alpha_0) \\ &\quad -a_{21}\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33}-\lambda & a_{34} \\ a_{42} & a_{43} & a_{44}-\lambda \end{vmatrix} - a_{31}\begin{vmatrix} a_{12} & a_{14} & a_{13} \\ a_{22}-\lambda & a_{24} & a_{23} \\ a_{42} & a_{44}-\lambda & a_{43} \end{vmatrix} \\ &\quad -a_{41}\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22}-\lambda & a_{23} & a_{24} \\ a_{32} & a_{33}-\lambda & a_{34} \end{vmatrix}\end{aligned}$$

Por los dos incisos anteriores,[24] es claro que al desarrollar los determinantes de los últimos tres términos se obtienen polinomios de grado 2. Así que

$$\det(A-\lambda I_4)=\lambda^4-(a_{11}+a_{33}+a_{44}+a_{22})\,\lambda^3+c_2\lambda^2+c_1\lambda+c_0$$

con $c_2, c_1, c_0 \in \mathbb{R}$.

Hemos hecho plausible la demostración, que puede hacerse por inducción, del teorema 5.29; la cual se deja como ejercicio al lector.

Teorema 5.29 *Sea A una matriz de orden n, entonces*

$$\det(A-\lambda I_n)=(-1)^n\lambda^n+c_{n-1}\lambda^{n-1}+\cdots+c_1\lambda+c_0,$$

donde los $c_i \in \mathbb{R}$*; es decir,* $\det(A-\lambda I_n)$ *es un* ***polinomio de grado*** *n. Además el coeficiente del término principal (el coeficiente de* λ^n*) es* $(-1)^n$ *y*[25] $c_{n-1}=(-1)^{n-1}\operatorname{tra}(A)$.

Definición 5.20 *Sea A una matriz cuadrada de orden n e* I_n *la matriz identidad de orden n.*

1. *Se denota y define el* ***polinomio característico*** *de la matriz A como*
$$p_A(\lambda)=\det(A-\lambda I_n).$$
2. *A la ecuación*
$$p_A(\lambda)=0;$$
esto es, $\det(A-\lambda I_n)=0$, *se le llama* ***ecuación característica*** *de la matriz A.*

Con esta terminología **los valores propios de** A **son las raíces reales del polinomio característico de** A; es decir, las soluciones reales de la ecuación característica. Vimos, en el teorema 5.26, que dos matrices similares tienen los mismos valores propios; ahora veamos que también tienen el mismo polinomio característico.

Teorema 5.30 *Sean A y B un par de matrices cuadradas de orden n que son similares. Entonces A y B tienen el mismo polinomio característico.*

DEMOSTRACIÓN ■ Sea P una matriz cuadrada invertible de orden n tal que $A=P^{-1}BP$. Entonces,

$$\begin{aligned}A-\lambda I_n&=P^{-1}BP-\lambda I_n\\&=P^{-1}BP-\lambda P^{-1}P\\&=P^{-1}(B-\lambda I_n)P\end{aligned}$$

[24] Note que intercambiamos dos columnas del determinante que multiplica a_{31}.

[25] Cfr. definición 4.2, página 242.

y por tanto

$$\begin{aligned}\det(A-\lambda I_n) &= \det(P^{-1}(B-\lambda I_n)P)\\ &= \det(P^{-1})\det(B-\lambda I_n)\det(P)\\ &= (\det(P))^{-1}\det(B-\lambda I_n)\det(P)\\ &= \det(B-\lambda I_n);\end{aligned}$$

esto es,

$$p_A(\lambda) = p_B(\lambda).\quad \blacksquare$$

Como consecuencia inmediata del teorema anterior tenemos que cualquier par de representaciones matriciales de un mismo operador lineal en un espacio de dimensión finita, por ser matrices similares, tiene el mismo polinomio característico.

Definición 5.21 *Si* $\mathbf{E}$ *es un espacio vectorial de dimensión finita y* $T:\mathbf{E}\to\mathbf{E}$ *es un operador lineal, se define el* ***polinomio característico*** *de* T, $p_T(\lambda)$, *como el polinomio característico de cualquier representación matricial de* T.

De lo precedente se concluye que una matriz cuadrada tiene a lo más n valores propios, pues un polinomio de grado n tiene a lo más n raíces reales. Lo mismo ocurre obviamente para un operador lineal en un espacio de dimensión finita n.

Teorema 5.31

1. *Si* A *es una matriz cuadrada de orden* n*, entonces* A *tiene a lo más* n *valores propios.*
2. *Si* $T:\mathbf{E}\to\mathbf{E}$ *es un operador lineal y* $\dim(\mathbf{E}) = n$*, entonces* T *tiene a lo más* n *valores propios.*

Método sintetizado para encontrar los valores y vectores propios correspondientes de una matriz

Resumimos a continuación la información precedente para calcular los valores y vectores propios correspondientes de una matriz cuadrada A de orden n:

1. Se calculan las soluciones reales, λ_i, de la ecuación característica

$$\det(A-\lambda I_n);$$

   es decir, las raíces reales del polinomio característico $p_A(\lambda)$.

2. Para cada λ_i se encuentran las soluciones no triviales del sistema no homogéneo

$$(A-\lambda_i I_n)\vec{x} = \vec{0}_{\mathbb{R}^n}$$

   y dichas soluciones serán los vectores propios correspondientes a λ_i.

❍ **Nota 5.10**

1. Dado que un polinomio de grado n puede tener raíces múltiples[26] (o repetidas), es obvio que el paso 2 únicamente se lleva a cabo para los valores propios λ_i distintos entre sí; es decir, para las raíces simples del polinomio característico.
2. Si $\mathbf{E}$ es un espacio de dimensión finita y $T:\mathbf{E}\to\mathbf{E}$ es un operador lineal, para calcular los valores propios y vectores propios correspondientes de T se aplica el procedimiento precedente a cualquier representación matricial de T.

▶ **Ejemplo 5.48** Sea $A=\begin{bmatrix} 2 & 1 & 0\\ -1 & 0 & 1\\ 1 & 3 & 1\end{bmatrix}$.

1. Encontrar los valores propios de A.
2. Hallar los vectores y espacios propios correspondientes.
3. Calcular las dimensiones de los espacios propios.◀

Solución Primero encontremos las raíces del polinomio característico. Para ello desarrollamos por cofactores $|A-\lambda I_3|$ por la tercera columna:

$$\begin{aligned}
p_A(\lambda) &= |A-\lambda I_3|\\
&= \begin{vmatrix} 2-\lambda & 1 & 0\\ -1 & -\lambda & 1\\ 1 & 3 & 1-\lambda\end{vmatrix}\\
&= -\begin{vmatrix} 2-\lambda & 1\\ 1 & 3\end{vmatrix} + (1-\lambda)\begin{vmatrix} 2-\lambda & 1\\ -1 & -\lambda\end{vmatrix}\\
&= -(3(2-\lambda)-1)+(1-\lambda)(-\lambda(2-\lambda))+1)\\
&= -(5-3\lambda)+(1-\lambda)(\lambda^2-2\lambda+1)\\
&= -(\lambda-1)^3-(5-3\lambda).
\end{aligned}$$

Sin embargo, este resultado no produce una factorización adecuada para calcular las raíces. Pero si desarrollamos el determinante por la primera fila obtenemos

$$\begin{aligned}
p_A(\lambda) &= (2-\lambda)\begin{vmatrix} -\lambda & 1\\ 3 & 1-\lambda\end{vmatrix} - \begin{vmatrix} -1 & 1\\ 1 & 1-\lambda\end{vmatrix}\\
&= (2-\lambda)[-\lambda(1-\lambda)-3]-(\lambda-1-1)\\
&= (2-\lambda)[\lambda^2-\lambda-3]-(\lambda-2)\\
&= (2-\lambda)[\lambda^2-\lambda-3]+(2-\lambda)\\
&= (2-\lambda)(\lambda^2-\lambda-3+1)\\
&= (2-\lambda)(\lambda^2-\lambda-2)\\
&= (2-\lambda)(\lambda-2)(\lambda+1)\\
&= -(\lambda-2)(\lambda-2)(\lambda+1)\\
&= -(\lambda-2)^2(\lambda+1).
\end{aligned}$$

[26]Por ejemplo, el polinomio de grado tres $p(\lambda)=(1-\lambda)^2(\lambda+2)$ tiene las raíces distintas $\lambda=1$ (raíz de multiplicidad 2) y la raíz simple $\lambda=-2$.

Así que las raíces de $p_A(\lambda)$ son $\lambda = 2$ (raíz doble) y $\lambda = -1$ (raíz simple).

1. Valores propios distintos: $\lambda_1 = 2$ y $\lambda_2 = -1$.
2. Los valores propios de λ_i, $i = 1, 2$, son las soluciones del sistema homogéneo $(A - \lambda_i I_3)\vec{v} = \vec{0}_{\mathbb{R}^3}$.

Para $\lambda_1 = 2$:

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & -2 & 1 \\ 1 & 3 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -1 \\ -1 & -2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

$$\therefore \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} r \\ 0 \\ r \end{bmatrix} = r\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Luego el conjunto de vectores propios para $\lambda_1 = 2$ es

$$\{(r, 0, r) \mid r \neq 0\}.$$

Para $\lambda_2 = -1$:

$$\begin{bmatrix} 3 & 1 & 0 \\ -1 & 1 & 1 \\ 1 & 3 & 2 \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 0 \\ 0 & 4 & 3 \\ 0 & 8 & 6 \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 0 \\ 0 & 4 & 3 \\ 0 & 0 & 0 \end{bmatrix};$$

$$\therefore \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} s \\ -3r \\ 4r \end{bmatrix} = r\begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}$$

y el conjunto de vectores propios de $\lambda_2 = -1$ es

$$\{(r, -3r, 4r) \mid r \neq 0\}.$$

3. Por lo precedente

$$E_2 = \operatorname{gn}((1, 0, 1)) \quad \text{y} \quad E_{-1} = \operatorname{gn}((1, -3, 4)).$$

Por tanto

$$\begin{aligned} \dim(E_2) &= 1, \\ \dim(E_{-1}) &= 1. \end{aligned}$$ ✔

En el ejemplo precedente $p_A = -(\lambda-1)^3 - (5-3\lambda)$ no tiene una forma adecuada para factorizar y calcular las raíces. Pero si desarrollamos, obtenemos

$$p_A(\lambda) = -\lambda^3 + 3\lambda^2 - 4$$

y la ecuación característica se puede escribir entonces como

$$\lambda^3 - 3\lambda^2 + 4 = 0.$$

Las raíces se pueden hallar mediante división sintética. Recuerde que en un polinomio que tiene coeficientes enteros y es mónico (el coeficiente de la mayor potencia es 1) si una raíz es entera, entonces debe dividir exactamente al término independiente, en este caso 4; y que si r es raíz del polinomio, entonces el residuo de dividir el polinomio entre $\lambda - r$ es cero. Para este caso, los divisores de 4 son $\pm1, \pm2, \pm4$. Al hacer división sintética se obtiene

$$\begin{array}{cccc|c} 1 & -3 & 0 & 4 & \\ & 2 & -2 & -4 & 2 \\ \hline 1 & -1 & -2 & 0 & \end{array}.$$

Los coeficientes del cociente son los primeros tres números de la última fila y el residuo está dado en la última columna de esta fila (0); por tanto, $\lambda^3 - 3\lambda^2 + 4 = (\lambda-2)(\lambda^2-\lambda-2)$; esto es, $\lambda^3 - 3\lambda^2 + 4 = (\lambda-2)(\lambda-2)(\lambda+1) = (\lambda-2)^2(\lambda+1)$. Las raíces son entonces $\lambda = 2$ (doble o de multiplicidad 2) y $\lambda = 1$ (simple o de multiplicidad 1), que son las mismas que encontramos antes.

5.3.2 Diagonalización

Al inicio de esta sección planteamos la posibilidad de encontrar una representación diagonal de un operador lineal T en un espacio de dimensión finita n; y en el teorema 5.25 (cfr. pág. 458) mostramos que las condiciones necesarias y suficientes para este fin son que existan n escalares reales λ_j y n vectores L.I. $\vec{e}_j$, $j = 1,\ldots,n$, tales que $T(\vec{e}_j) = \lambda_j\vec{e}_j\ \forall j$. Esto es, que T tenga n valores propios λ_j con sendos vectores propios correspondientes $\vec{e}_j$ que sean linealmente independientes; es decir, que formen una base del espacio. En tal caso dicha representación diagonal es $\operatorname{diag}(\lambda_1,\ldots,\lambda_n)$. Si A es la representación matricial de T relativa a una base $\mathscr{B}$, los valores propios de T y de A son los mismos; por tanto, que T tenga una representación matricial diagonal equivale a que exista una matriz C tal que $D = C^{-1}AC$, donde $D = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$; pues las representaciones matriciales del operador T son matrices similares. Entonces el objetivo se limita nuevamente a trabajar con matrices cuadradas y determinar bajo qué condiciones se puede llevar a efecto esta "diagonalización".

Definición 5.22 *Sea $A \in \mathfrak{M}_n$ una matriz cuadrada real de orden n. A es **diagonalizable** si existe un par de matrices $C, D \in \mathfrak{M}_n$, con D una matriz diagonal y C una matriz invertible, tales que*

$$D = C^{-1}AC.$$

Se dice entonces que el par (C, D) es una diagonalización *para la matriz A.*

Sea A una matriz diagonalizable de orden n, y (C,D) una diagonalización para A, con $D = \text{diag}(\lambda_1,\ldots,\lambda_n)$. Puesto que A y D son similares tienen los mismos valores propios (cfr. teorema 5.26); de ahí que los valores propios de A son[27] $\lambda_1,\ldots,\lambda_n$. Sean $\vec{K}_1,\ldots,\vec{K}_n$ las columnas de la matriz C; entonces, puesto que $AC = CD$, se tiene, de (1.2), página 11,

$$[A\vec{K}_1\ A\vec{K}_2\ \cdots\ A\vec{K}_n] = [\lambda_1\vec{K}_1\ \lambda_2\vec{K}_2\ \cdots\ \lambda_n\vec{K}_n];$$

por lo que $A\vec{K}_i = \lambda_i\vec{K}_i$ y por ende $\vec{K}_i$ es un vector propio de la matriz A correspondiente al valor propio λ_i para cada $i = 1,\ldots,n$. De esto y del teorema 5.25, con $T = T_A$, se demuestra el siguiente teorema.

Teorema 5.32 *Sea A una matriz cuadrada de orden n. Entonces A es diagonalizable si y sólo si A tiene n valores propios $\lambda_j \in \mathbb{R}$ con vectores propios correspondientes $\vec{u}_j$, de sendos valores propios, linealmente independientes; es decir, los vectores propios $\vec{u}_1,\ldots,\vec{u}_n$ forman una base de $\mathbb{R}^n$. En tal caso una diagonalización para A está dada por el par (C,D); donde C es la matriz cuyas columnas son los vectores $\vec{u}_j$, $j = 1,2,\ldots,n$, y $D = \text{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$.*

Necesitaremos del siguiente lema para tener una consecuencia inmediata del teorema anterior y poder probar el principal resultado de esta sección (el teorema 5.33).

Lema 5.2

1. *Sean T un operador lineal definido en un espacio vectorial y $\lambda_1,\ldots,\lambda_k$ valores propios de T distintos entre sí. Supongamos que $\{\vec{u}_{11},\ldots,\vec{u}_{1m_1}\}$ es un conjunto de vectores propios L.I correspondientes a λ_1, $\{\vec{u}_{21},\ldots,\vec{u}_{2m_2}\}$ es un conjunto de vectores propios L.I. correspondientes a λ_2, etc., y $\{\vec{u}_{k1},\ldots,\vec{u}_{km_k}\}$ es un conjunto de vectores propios L.I. correspondientes a λ_k. Entonces los vectores*

$$\vec{u}_{11},\ldots,\vec{u}_{1m_1},\vec{u}_{21},\ldots,\vec{u}_{2m_2},\ldots,\vec{u}_{k1},\ldots,\vec{u}_{km_k} \tag{5.23}$$

son linealmente independientes.

2. *Sean A una matriz cuadrada de orden n y $\lambda_1,\ldots,\lambda_k$ valores propios de A distintos entre sí. Sean $\{\vec{u}_{j1},\ldots,\vec{u}_{jm_j}\}$ conjuntos L.I. de vectores propios correspondientes a λ_j para cada $j = 1,\ldots,k$. Entonces los vectores*

$$\vec{u}_{11},\ldots,\vec{u}_{1m_1},\vec{u}_{21},\ldots,\vec{u}_{2m_2},\ldots,\vec{u}_{k1},\ldots,\vec{u}_{km_k}$$

son linealmente independientes.

DEMOSTRACIÓN ■ 1. Procedamos por inducción sobre k. Si $k = 1$, por hipótesis $\{\vec{u}_{11},\ldots,\vec{u}_{1m_1}\}$ es L.I. Sea k un entero mayor a 1 y supongamos que el resultado es cierto para todo conjunto de $k-1$ valores propios de T distintos entre sí. Si uno de los vectores en (5.23) es combinación lineal de los demás, digamos $\vec{u}_{km_k}$ (de ser necesario podemos renombrar estos vectores y los valores propios para que esto suceda así), entonces existen escalares

$$\alpha_{11},\ldots,\alpha_{1m_1},\alpha_{21},\ldots,\alpha_{2m_2},\ldots,\alpha_{k1},\ldots,\alpha_{km_k-1}$$

[27] Es evidente que los valores propios de una matriz diagonal $D = \text{diag}(\lambda_1,\ldots,\lambda_n)$ son los elementos λ_i de su diagonal.

tales que

$$\vec{u}_{km_k} = \sum_{i=1}^{m_1} \alpha_{1i}\vec{u}_{1i} + \sum_{i=1}^{m_2} \alpha_{2i}\vec{u}_{2i} + \cdots + \sum_{i=1}^{m_k-1} \alpha_{ki}\vec{u}_{ki} \tag{5.24}$$

Por tanto

$$T(\vec{u}_{km_k}) = \sum_{i=1}^{m_1} \alpha_{1i}T(\vec{u}_{1i}) + \sum_{i=1}^{m_2} \alpha_{2i}T(\vec{u}_{2i}) + \cdots + \sum_{i=1}^{m_k-1} \alpha_{ki}T(\vec{u}_{ki})$$

y entonces

$$\lambda_k\vec{u}_{km_k} = \sum_{i=1}^{m_1} \alpha_{1i}\lambda_1\vec{u}_{1i} + \sum_{i=1}^{m_2} \alpha_{2i}\lambda_2\vec{u}_{2i} + \cdots + \sum_{i=1}^{m_k-1} \alpha_{ki}\lambda_k\vec{u}_{ki} \tag{5.25}$$

Pero, por (5.24),

$$\lambda_k\vec{u}_{km_k} = \sum_{i=1}^{m_1} \alpha_{1i}\lambda_k\vec{u}_{1i} + \sum_{i=1}^{m_2} \alpha_{2i}\lambda_k\vec{u}_{2i} + \cdots + \sum_{i=1}^{m_k-1} \alpha_{ki}\lambda_k\vec{u}_{ki} \tag{5.26}$$

De (5.25) y (5.26) se desprende

$$\sum_{i=1}^{m_1} \alpha_{1i}(\lambda_1 - \lambda_k)\vec{u}_{1i} + \cdots + \sum_{i=1}^{m_{k-1}} \alpha_{(k-1)i}(\lambda_{k-1} - \lambda_k)\vec{u}_{(k-1)i} = \vec{0}_{\mathbf{E}} \tag{5.27}$$

Por hipótesis de inducción, ya que los λ_j, $j = 1, \ldots, k-1$, son distintos entre sí, se tiene que los vectores

$$\vec{u}_{11}, \ldots, \vec{u}_{1m_1}, \vec{u}_{21}, \ldots, \vec{u}_{2m_2}, \ldots, \vec{u}_{(k-1)1}, \ldots, \vec{u}_{(k-1)m_{k-1}}$$

son L.I.; así que (5.27) implica

$$\begin{aligned}
\alpha_{11}(\lambda_1 - \lambda_k) &= \cdots = & \alpha_{1m_1}(\lambda_1 - \lambda_k) &= 0\\
\alpha_{21}(\lambda_2 - \lambda_k) &= \cdots = & \alpha_{2m_2}(\lambda_2 - \lambda_k) &= 0\\
&\vdots & &\\
\alpha_{(k-1)1}(\lambda_{k-1} - \lambda_k) &= \cdots = & \alpha_{(k-1)m_{k-1}}(\lambda_{k-1} - \lambda_k) &= 0
\end{aligned}$$

y puesto que $\lambda_j - \lambda_k \neq 0$ para todo $j = 1, 2, \ldots, k-1$, se concluye que

$$\begin{aligned}
\alpha_{11} &= \cdots = & \alpha_{1m_1} &= 0\\
\alpha_{21} &= \cdots = & \alpha_{2m_2} &= 0\\
&\vdots & &\\
\alpha_{(k-1)1} &= \cdots = & \alpha_{(k-1)m_{k-1}} &= 0.
\end{aligned}$$

Al sustituir estos α_{ij} en (5.24) obtenemos

$$\vec{u}_{km_k} = \sum_{i=1}^{m_k-1} \alpha_{ki}\vec{u}_{ki}$$

que es imposible, pues por hipótesis los vectores $\vec{u}_{k1}, \ldots, \vec{u}_{1m_k}$ son L.I. Luego, los vectores en (5.23) tienen que ser independientes.

2. Es consecuencia inmediata del inciso anterior al tomar $T = T_A$. ■

Corolario 5.2 *Si A es una matriz cuadrada de orden n que tiene n valores propios $\lambda_1, \ldots, \lambda_n$ distintos entre sí, entonces A es diagonalizable y una diagonalización para A es el par (C, D); donde C es la matriz cuya columna j, $j = 1, \ldots, n$, es un vector propio $\vec{u}_j$ correspondiente al valor propio λ_j y $D = diag(\lambda_1, \ldots, \lambda_n)$.*

DEMOSTRACIÓN ■ Por el lema 5.2 los vectores propios $\vec{u}_j$, $j = 1, \ldots, n$, son L.I. y por ende $\{\vec{u}_1, \ldots, \vec{u}_n\}$ es una base de $\mathbb{R}^n$; y la conclusión es entonces consecuencia inmediata del teorema 5.32. ■

► **Ejemplo 5.49** Sea $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Entonces

$$\begin{aligned} p_A(\lambda) &= \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)^2 - 4 \\ &= \lambda^2 - 2\lambda - 3 \\ &= (\lambda - 3)(\lambda + 1); \end{aligned}$$

así que los valores propios de A son $\lambda_1 = 3$ y $\lambda_2 = -1$, $n = 2$ valores propios distintos. Por el corolario 5.2, A es diagonalizable. Para hallar una diagonalización necesitamos calcular vectores propios.

Para λ_1:

$$\begin{aligned} \begin{bmatrix} 1-\lambda_1 & 2 \\ 2 & 1-\lambda_1 \end{bmatrix} &= \begin{bmatrix} 1-3 & 2 \\ 2 & 1-3 \end{bmatrix} \\ &= \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

y por tanto los vectores propios correspondientes tienen la forma

$$r \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad r \neq 0.$$

Para λ_2:

$$\begin{aligned} \begin{bmatrix} 1-\lambda_2 & 2 \\ 2 & 1-\lambda_2 \end{bmatrix} &= \begin{bmatrix} 1-(-1) & 2 \\ 2 & 1-(-1) \end{bmatrix} \\ &= \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

y entonces los vectores propios correspondientes tienen la forma

$$r \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad r \neq 0.$$

Así,

$$C = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, D = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$$

es una diagonalización para A.

Comprobación:

$$CD = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 3 & 1 \end{bmatrix}$$

y

$$AC = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 3 & 1 \end{bmatrix}$$

Por tanto

$$AC = CD$$

que equivale a

$$D = C^{-1}AC. ◀$$

El corolario 5.2 únicamente establece condiciones suficientes para que una matriz sea diagonalizable; es decir, puede ser que una matriz sea diagonalizable aunque su polinomio característico tenga raíces múltiples (cfr. ejemplo 5.50). De hecho, vamos a establecer condiciones necesarias y suficientes para que una matriz sea diagonalizable (teorema 5.33) que cubrirán todos esos casos. Antes de ello necesitamos un poco más de terminología.

Definición 5.23

1. *Sean* $\mathbf{E}$ *un espacio vectorial de dimensión finita* n*,* T *un operador lineal en* $\mathbf{E}$ *y* λ *un valor propio de* T*.*
   (a) *A la multiplicidad de* λ *como raíz del polinomio característico de* T *se le llama la* ***multiplicidad algebraica*** *de este valor propio y la denotaremos, en este libro, por* $\mu_a(\lambda)$*.*
   (b) *A la dimensión del espacio propio* E_λ *se le dice* ***multiplicidad geométrica*** *de* λ *y la representaremos, en este texto, mediante el símbolo* $\mu_g(\lambda)$*.*
2. *Sea* A *una matriz de orden* n *y* λ *un valor propio de* A*.*
   (a) *Se llama* ***multiplicidad algebraica*** *de* λ *a la multiplicidad de este valor propio como raíz del polinomio característico de la matriz* A *y se le representa por el símbolo* $\mu_a(\lambda)$*.*
   (b) *Se define la* ***multiplicidad geométrica*** *de* λ *como la dimensión del espacio propio* E_λ*; esto es,* $\mu_g(\lambda) = \dim(E_\lambda)$*.*

❍ **Nota 5.11**

1. Sabemos, del teorema 5.12, que si $\vec{u}_1, \ldots, \vec{u}_k$ son vectores de un espacio vectorial $\mathbf{E}$ de dimensión finita n, $\mathscr{B}$ es una base de este espacio y $[\vec{u}_i]_{\mathscr{B}}$, $i = 1, \ldots, k$, son los respectivos vectores de coordenadas relativos a dicha base, entonces los vectores $\vec{u}_i$ son L.I. en $\mathbf{E}$ si y sólo si los vectores $[\vec{u}_i]_{\mathscr{B}}$ son L.I. en $\mathbb{R}^n$. De esto se desprende que la multiplicidad geométrica de un valor propio λ de un operador lineal $T : \mathbf{E} \to \mathbf{E}$ es la misma que la multiplicidad geométrica de este valor propio para cualquier representación matricial de T.
2. Si A es una matriz cuadrada, es evidente que los conceptos de multiplicidad algebraica y geométrica son exactamente los mismos que los correspondientes conceptos para el operador T_A.

Un resultado preliminar para lograr el objetivo final de dar condiciones necesarias y suficientes para la diagonalización de una matriz está contenido en el lema 5.3.

Lema 5.3

1. *Sean T un operador lineal en un espacio vectorial de dimensión n y λ_0 un valor propio de este operador, entonces*

$$\mu_g(\lambda_0) \leq \mu_a(\lambda_0). \qquad (5.28)$$

2. *Sea A una matriz de orden n. Entonces, para cada valor propio λ_0 de A se tiene (5.28).*

DEMOSTRACIÓN 1. Sean $m = \mu_g(\lambda_0)$ y $\mathscr{B} = \{\vec{e}_1, \ldots, \vec{e}_m\}$ una base del subespacio propio E_{λ_0}. Completemos $\mathscr{B}$ a una base $\{\vec{e}_1, \ldots, \vec{e}_m, \vec{e}_{m+1}, \ldots, \vec{e}_n\}$ de $\mathbf{E}$ mediante el procedimiento dado en la página 167. Entonces

$$T(\vec{e}_i) = \lambda_0 \vec{e}_i \quad \forall i = 1, \ldots, m.$$

De aquí se sigue que la representación matricial de T relativa a esta base tiene la forma

$$[T]_{\mathscr{B}} = \begin{bmatrix} D & B_1 \\ \mathscr{O} & B_2 \end{bmatrix},$$

donde $D = \operatorname{diag}(\underbrace{\lambda_0, \ldots, \lambda_0}_{m})$, B_1 y B_2 son matrices de tamaños $m \times (n-m)$ y $(n-m) \times (n-m)$, respectivamente y $\mathscr{O}$ es la matriz cero de tamaño $(n-m) \times m$. Entonces

$$p_T(\lambda) = \begin{vmatrix} D - \lambda I_m & B_1 \\ \mathscr{O} & B_2 - \lambda I_{n-m} \end{vmatrix}.$$

Es fácil mostrar que

$$\begin{vmatrix} D - \lambda I_m & B_1 \\ \mathscr{O} & B_2 - \lambda I_{n-m} \end{vmatrix} = \det(D - \lambda I_m)\det(B_2 - \lambda I_{n-m});$$

por tanto,

$$p_T(\lambda) = (\lambda_0 - \lambda)^m \det(B_2 - \lambda I_{n-m}).$$

De aquí que λ_0 es raíz del polinomio característico de multiplicidad al menos m; esto es,

$$\mu_g(\lambda_0) \leq \mu_a(\lambda_0).$$

2. Es inmediata al aplicar el primer inciso a $T = T_A$. ■

Supongamos ahora que A es una matriz cuadrada de orden n con n valores propios tal que para cada uno de éstos la multiplicidad algebraica coincide con la multiplicidad geométrica. Esto es, si A tiene k valores propios distintos entre sí, $\lambda_1, \ldots, \lambda_k$, entonces $\mu_g(\lambda_i) = \mu_a(\lambda_i) = m_i$, para cada $i = 1, \ldots, k$. Sean entonces $\vec{u}_1^j, \ldots, \vec{u}_{m_j}^j$ vectores propios L.I. correspondientes al valor propio λ_j, para cada $j = 1, \ldots, k$. En virtud del lema 5.2 inciso 2, ya que los valores propios $\lambda_1, \ldots, \lambda_k$ son distintos entre sí, el conjunto

$$\{\vec{u}_1^1, \ldots, \vec{u}_{m_1}^1, \vec{u}_1^2, \ldots, \vec{u}_{m_2}^2, \ldots, \vec{u}_1^k, \ldots, \vec{u}_{m_k}^k\}$$

es linealmente independiente. Finalmente, ya que

$$m_1 + m_2 + \cdots + m_k = n$$

(pues la suma de las multiplicidades de las raíces distintas de un polinomio es igual al grado del mismo), del teorema 5.32, se concluye que A es diagonalizable con diagonalización la matriz C que tiene por columnas a los vectores $\vec{u}_1^1, \ldots, \vec{u}_{m_1}^1, \vec{u}_1^2, \ldots, \vec{u}_{m_2}^2, \ldots, \vec{u}_1^k, \ldots, \vec{u}_{m_k}^k$ y

$$D = \operatorname{diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{m_1}, \underbrace{\lambda_2, \ldots, \lambda_2}_{m_2}, \ldots, \underbrace{\lambda_k, \ldots, \lambda_k}_{m_k}).$$

Inversamente, supongamos que la matriz es diagonalizable. Sea (C, D) una diagonalización para esta matriz con $D = \operatorname{diag}(\beta_1, \beta_2, \ldots, \beta_n)$. Entonces, si $\lambda_1, \lambda_2, \ldots, \lambda_k$, son los valores distintos entre sí de la diagonal de D, se tiene

$$p_A(\lambda) = (\lambda_1 - \lambda)^{m_1}(\lambda_2 - \lambda)^{m_2} \cdots (\lambda_k - \lambda)^{m_k};$$

donde $m_i = \mu_a(\lambda_i)$ para cada $i = 1, \ldots, k$. De aquí, ya que las columnas de C son vectores L.I. por ser ésta una matriz invertible, existen por lo menos m_i vectores L.I. en E_{λ_i} para cada $i = 1, \ldots, k$; pues, para cada columna $\vec{K}_j$ de C, $A\vec{K}_j = \beta_j \vec{K}_j$ (cfr. la discusión que sigue de la definición 5.22); lo cual prueba que

$$\mu_a(\lambda_i) \le \mu_g(\lambda_i) \quad \text{para cada } i = 1, \ldots, k. \tag{5.29}$$

De (5.28) del lema 5.3 y (5.29) se desprende

$$\mu_a(\lambda_i) = \mu_g(\lambda_i)$$

para cada valor propio de A. Hemos probado así el resultado que da condiciones necesarias y suficientes para que una matriz sea diagonalizable y que hacemos patente en el teorema 5.33.

Condiciones necesarias y suficientes para que una matriz sea diagonalizable

Teorema 5.33 *Sea A una matriz cuadrada de orden n que tiene n valores propios. Entonces A es diagonalizable si y sólo si para cada valor propio de A la multiplicidad algebraica coincide con la multiplicidad geométrica; esto es,*

$$\mu_a(\lambda) = \mu_g(\lambda)$$

para cada valor propio λ de A. En este caso, si $\lambda_1, \ldots, \lambda_k$ son los valores propios de A distintos entre sí con sendas multiplicidades algebraicas $\mu_a(\lambda_j) = m_j$, $j = 1, \ldots, k$, y $\vec{u}_1^j, \ldots, \vec{u}_{m_j}^j$ son vectores L.I. correspondientes a cada valor propio λ_j, entonces una diagonalización para A es el par (C, D), donde C es la matriz que tiene por columnas a los vectores

$$\vec{u}_1^1, \ldots, \vec{u}_{m_1}^1, \vec{u}_1^2, \ldots, \vec{u}_{m_2}^2, \ldots, \vec{u}_1^k, \ldots, \vec{u}_{m_k}^k$$

y

$$D = \operatorname{diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{m_1}, \underbrace{\lambda_2, \ldots, \lambda_2}_{m_2}, \ldots, \underbrace{\lambda_k, \ldots, \lambda_k}_{m_k}).$$

▶ **Ejemplo 5.50** Determinar si la matriz

$$A = \begin{bmatrix} 1 & -3 & 3 \\ 0 & -5 & 6 \\ 0 & -3 & 4 \end{bmatrix}$$

es diagonalizable y encontrar, en su caso, una diagonalización para ella.◀

Solución

$$\begin{aligned} p_A(\lambda) &= \begin{vmatrix} 1-\lambda & -3 & 3 \\ 0 & -5-\lambda & 6 \\ 0 & -3 & 4-\lambda \end{vmatrix} \\ &= (1-\lambda)\begin{vmatrix} -5-\lambda & 6 \\ -3 & 4-\lambda \end{vmatrix} \\ &= (1-\lambda)(-(5-\lambda)(4-\lambda)+18) \\ &= (1-\lambda)(\lambda^2+\lambda-2) \\ &= (1-\lambda)(\lambda+2)(\lambda-1) \\ &= -(1-\lambda)^2(\lambda+2). \end{aligned}$$

Así que los vectores propios de A, distintos entre sí, son

$$\lambda_1 = 1 \text{ y } \lambda_2 = -2;$$

con multiplicidades algebraicas

$$\mu_a(\lambda_1) = 2 \text{ y } \mu_a(\lambda_2) = 1.$$

Para $\lambda_1 = 1$:

$$(A - \lambda_1 I_3) = \begin{bmatrix} 0 & -3 & 3 \\ 0 & -6 & 6 \\ 0 & -3 & 3 \end{bmatrix} \sim \begin{bmatrix} 0 & -3 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

luego

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} s \\ r \\ r \end{bmatrix} = s\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + r\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Por tanto, $\mu_g(\lambda_1) = \dim(E_1) = 2$.

Para $\lambda_2 = -2$:

$$(A + 2I_3) = \begin{bmatrix} 3 & -3 & 3 \\ 0 & -3 & 6 \\ 0 & -3 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & -3 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix};$$

luego

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} r \\ 2r \\ r \end{bmatrix} = r\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Por tanto, $\mu_g(\lambda_2) = \dim(E_{-2}) = 1$.

Puesto que $\mu_a(\lambda_1)=\mu_g(\lambda_1)$ y $\mu_a(\lambda_2)=\mu_g(\lambda_2)$, A es diagonalizable y una diagonalización para A es el par (C,D), donde

$$C=\begin{bmatrix}1&0&1\\0&1&2\\0&1&1\end{bmatrix}\text{ y }D=\begin{bmatrix}1&0&0\\0&1&0\\0&0&-2\end{bmatrix}$$

de acuerdo con el teorema 5.33. El lector puede comprobar, verificando la igualdad de los productos[28] AC y CD, que efectivamente $D=C^{-1}AC$. ✔

▶ **Ejemplo 5.51** Sea $A=\begin{bmatrix}2&1&0\\-1&0&1\\1&3&1\end{bmatrix}$. En el ejemplo 5.48 (cfr. pág. 469) vimos que los valores propios de A son $\lambda_1=2$ y $\lambda_2=-1$ con multiplicidades algebraicas $\mu_a(\lambda_1)=2$ y $\mu_a(\lambda_2)=1$; y que $\dim(E_2)=1$ y $\dim(E_{-1})=1$. Luego $\mu_a(\lambda_1)\neq\mu_g(\lambda_1)$; por tanto, A no es diagonalizable.◀

▶ **Ejemplo 5.52** Sea

$$A=\begin{bmatrix}1&3&0&0\\0&-2&0&0\\3&3&1&-3\\3&3&0&-2\end{bmatrix}.$$

Determinar si A es diagonalizable y, en caso afirmativo, encontrar una diagonalización para esta matriz.◀

Solución

$$\begin{aligned}p_A(\lambda)&=\begin{vmatrix}1-\lambda&3&0&0\\0&-2-\lambda&0&0\\3&3&1-\lambda&-3\\3&3&0&-2-\lambda\end{vmatrix}\\&=(1-\lambda)\begin{vmatrix}-2-\lambda&0&0\\3&1-\lambda&-3\\3&0&-2-\lambda\end{vmatrix}\\&=(1-\lambda)(-2-\lambda)\begin{vmatrix}1-\lambda&-3\\0&-2-\lambda\end{vmatrix}\\&=(1-\lambda)(-2-\lambda)(1-\lambda)(-2-\lambda)\\&=(1-\lambda)^2(2+\lambda)^2,\end{aligned}$$

por lo que los valores propios distintos de A son $\lambda_1=-2$ y $\lambda_2=1$, con multiplicidades algebraicas $\mu_a(\lambda_1)=2$ y $\mu_a(\lambda_2)=2$. Resolvamos los sistemas $(A-\lambda_iI_4)\vec{u}=\vec{0}_{\mathbb{R}^n}$ para $i=1,2$.

Para $\lambda_1=-2$:

$$\begin{aligned}A-(-2)I_4=\begin{bmatrix}3&3&0&0\\0&0&0&0\\3&3&3&-3\\3&3&0&0\end{bmatrix}&\sim\begin{bmatrix}1&1&1&-1\\1&1&0&0\\1&1&0&0\\0&0&0&0\end{bmatrix}\\&\sim\begin{bmatrix}1&1&1&-1\\0&0&-1&1\\0&0&-1&1\\0&0&0&0\end{bmatrix}\sim\begin{bmatrix}1&1&1&-1\\0&0&-1&1\\0&0&0&0\\0&0&0&0\end{bmatrix},\end{aligned}$$

[28] Observe que se comprueba la igualdad de los productos AC y CD para evitar el cálculo de C^{-1}.

por lo que $x_4 = x_3$, $x_1 = -x_2 - x_3 + x_4 = -x_2$; esto es, la solución del sistema homogéneo es

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -s \\ s \\ r \\ r \end{bmatrix} = s\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + r\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

y entonces $\dim(E_{-2}) = 2$.

Para $\lambda_2 = 1$:

$$\begin{aligned} A - I_4 &= \begin{bmatrix} 0 & 3 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 3 & 3 & 0 & -3 \\ 3 & 3 & 0 & -3 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

de donde $x_2 = 0$, $x_1 = x_4$; por tanto, la solución del sistema homogéneo es

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} r \\ 0 \\ s \\ r \end{bmatrix} = r\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + s\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

así que $\dim(E_1) = 2$.

Puesto que $\mu_a(\lambda_1) = 2 = \mu_g(\lambda_1)$ y $\mu_a(\lambda_2) = 2 = \mu_a(\lambda_2)$, A es diagonalizable y una diagonalización para esta matriz está dada por

$$C = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Comprobación:

$$\begin{aligned} AC &= \begin{bmatrix} 1 & 3 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 3 & 3 & 1 & -3 \\ 3 & 3 & 0 & -2 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 1 & 0 \\ -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 1 \\ 0 & -2 & 1 & 0 \end{bmatrix} \end{aligned}$$

y

$$CD = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 0 & 1 & 0 \\ -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 1 \\ 0 & -2 & 1 & 0 \end{bmatrix}. \checkmark$$

Propiedades de matrices diagonalizables

Sea A una matriz cuadrada de orden n que es diagonalizable, con diagonalización (C,D) y $D = \text{diag}(\lambda_1,\ldots,\lambda_n)$. Entonces $D = C^{-1}AC$, de donde

$$\det(A) = \det(D) = \lambda_1 \cdots \lambda_n;$$

esto es, el determinante de A es el producto de los valores propios de A. Por otro lado,

$$\begin{aligned} D^2 &= (C^{-1}AC)(C^{-1}AC) \\ &= C^{-1}A^2C, \end{aligned}$$

de donde se infiere, lo cual se puede probar por inducción, que

$$D^m = C^{-1}A^mC$$

para todo entero positivo m; y puesto que $D^m = \text{diag}(\lambda_1^m,\ldots,\lambda_n^m)$, se tiene

$$A^m = C\,\text{diag}(\lambda_1^m,\ldots,\lambda_n^m)C^{-1}.$$

Resumimos la información precedente en el siguiente teorema.

Teorema 5.34 *Si A es una matriz cuadrada de orden n que es diagonalizable con diagonalización (C,D), $D = diag(\lambda_1,\ldots,\lambda_n)$, entonces $\lambda_1,\ldots,\lambda_n$ son los valores propios de A y*

1. $$\det(A) = \lambda_1 \cdots \lambda_n. \tag{5.30}$$

2. *Para todo entero positivo m,*

$$A^m = C\,diag(\lambda_1^m,\ldots,\lambda_n^m)C^{-1}.$$

❍ **Nota 5.12** En realidad mostraremos, en el siguiente apartado, que la igualdad (5.30) es válida para toda matriz A aunque ella nos sea diagonalizable (cfr. teorema 5.36, pág. 486); esto es, que el determinante de toda matriz cuadrada es el producto de los valores propios de la misma.

5.3.3 Valores propios complejos y diagonalización sobre $\mathbb{C}$

Valores propios complejos, propiedades

Sabemos de educación elemental que las raíces de un polinomio no siempre son todas ellas números reales; de hecho, aun en casos muy simples, se obtienen raíces complejas. El campo $\mathbb{C}$ de números complejos *contiene*, algebraicamente, al campo $\mathbb{R}$ de los números reales mediante la inclusión $a \hookrightarrow a+0i$. Es entonces evidente que las matrices con entradas reales, $\mathfrak{M}_{m\times n}$, pueden incluirse algebraicamente en el conjunto de matrices con entradas complejas, $\mathfrak{M}_{m\times n}(\mathbb{C})$, mediante la identificación antes mencionada de los números reales a con los números complejos de la forma $a+0i$. De esta manera podemos extender los conceptos de valores y vectores propios a este tipo de matrices para considerar los casos en los que el polinomio característico tenga raíces complejas. Lejos de ser una simple generalización esta idea resulta sumamente útil tanto en aspectos teóricos como en aplicaciones y su uso, a estas alturas, es inevitable.[29]

Definición 5.24 *(Valores propios complejos de una matriz) Sea A una matriz cuadrada de orden n con entradas complejas, i.e., $A \in \mathfrak{M}_n(\mathbb{C})$. Se dice que $\lambda \in \mathbb{C}$ es un valor propio de A si existe un vector $\vec{u} \in \mathbb{C}^n$, $\vec{u} \neq (\underbrace{0,0,\ldots,0}_{n})$, tal que $A\vec{u} = \lambda\vec{u}$. En este caso, a $\vec{u}$ se le llama vector propio de A correspondiente a λ.*

Todo lo que hemos desarrollado, relativo a valores y vectores propios de matrices cuadradas con componentes reales, sigue siendo válido para matrices cuadradas con entradas complejas simplemente considerando la definición 5.24 en lugar del caso particular de ésta contenido en la definición 5.19.

▶ **Ejemplo 5.53** Sea $A = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$. Encontrar los valores y vectores propios correspondientes de la matriz A.◀

$$\begin{aligned} \det(A-\lambda I_2) &= \begin{vmatrix} 1-\lambda & -1 \\ 1 & 2-\lambda \end{vmatrix} \\ &= (1-\lambda)(2-\lambda)+1 \\ &= 2-3\lambda+\lambda^2+1 \\ &= \lambda^2-3\lambda+3. \end{aligned}$$

Las raíces del polinomio característico son entonces

$$\begin{aligned} \lambda &= \frac{3\pm\sqrt{9-12}}{2} \\ &= \frac{3}{2}\pm\frac{\sqrt{-3}}{2} \\ &= \frac{3}{2}\pm\frac{\sqrt{3i^2}}{2} \\ &= \frac{3}{2}\pm\frac{\sqrt{3}}{2}i. \end{aligned}$$

[29] El lector que hasta aquí ha omitido las subsecciones 1.1.5, 1.2.8, 2.1.5, 2.2.5, 3.5 y la sección B.1 del apéndice B, éste es el momento para invertir un poco de tiempo en su estudio; el esfuerzo no será demasiado y bien valdrá la pena.

Así, los valores propios de A son $\lambda_1 = \frac{3}{2} + \frac{\sqrt{3}}{2}i$ y $\lambda_2 = \frac{3}{2} - \frac{\sqrt{3}}{2}i$. Resolvamos los sistemas homogéneos $(A - \lambda_j I_2)\vec{u} = \vec{0}_{\mathbb{C}} = [\ 0 \quad 0\ ]^t$:

- Para $\lambda_1 = \frac{3}{2} + \frac{\sqrt{3}}{2}i$:

$$\begin{bmatrix} 1 - \frac{3}{2} - \frac{\sqrt{3}}{2}i & -1 \\ 1 & 2 - \frac{3}{2} - \frac{\sqrt{3}}{2}i \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} - \frac{\sqrt{3}}{2}i & -1 \\ 1 & \frac{1}{2} - \frac{\sqrt{3}}{2}i \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & \frac{1}{2} - \frac{\sqrt{3}}{2}i \\ -\frac{1}{2} - \frac{\sqrt{3}}{2}i & -1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & \frac{1}{2} - \frac{\sqrt{3}}{2}i \\ 0 & 0 \end{bmatrix}$$

y por tanto,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) r \\ r \end{bmatrix}$$
$$= r \begin{bmatrix} -\frac{1}{2} + \frac{\sqrt{3}}{2}i \\ 1 \end{bmatrix}.$$

Luego, los vectores propios correspondientes a λ_1 son todos los vectores en $\mathbb{C}^2$ de la forma

$$r\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i, 1\right)$$

con r cualquier número complejo distinto de cero.

- Para $\lambda_2 = \frac{3}{2} - \frac{\sqrt{3}}{2}i$:

$$\begin{bmatrix} 1 - \frac{3}{2} + \frac{\sqrt{3}}{2}i & -1 \\ 1 & 2 - \frac{3}{2} + \frac{\sqrt{3}}{2}i \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} + \frac{\sqrt{3}}{2}i & -1 \\ 1 & \frac{1}{2} + \frac{\sqrt{3}}{2}i \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & \frac{1}{2} + \frac{\sqrt{3}}{2}i \\ -\frac{1}{2} + \frac{\sqrt{3}}{2}i & -1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & \frac{1}{2} + \frac{\sqrt{3}}{2}i \\ 0 & 0 \end{bmatrix}$$

por lo que

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -\left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) r \\ r \end{bmatrix}$$
$$= s \begin{bmatrix} \frac{1}{2} + \frac{\sqrt{3}}{2}i \\ -1 \end{bmatrix}.$$

Entonces los vectores propios correspondientes a λ_2 tienen la forma $s\left(\frac{1}{2} + \frac{\sqrt{3}}{2}i, -1\right)$ con $s \in \mathbb{C} - \{0\}$.

Comprobación: Para λ_1, $\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} + \frac{\sqrt{3}}{2}i \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} + \frac{\sqrt{3}}{2}i \\ \frac{3}{2} + \frac{\sqrt{3}}{2}i \end{bmatrix}$ y

$$\left(\frac{3}{2} + \frac{\sqrt{3}}{2}i\right) \begin{bmatrix} -\frac{1}{2} + \frac{\sqrt{3}}{2}i \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} + \frac{\sqrt{3}}{2}i \\ \frac{3}{2} + \frac{\sqrt{3}}{2}i \end{bmatrix}.$$

Para λ_2, $\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{2}+\frac{\sqrt{3}}{2}i \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{3}{2}+\frac{\sqrt{3}}{2}i \\ -\frac{3}{2}+\frac{\sqrt{3}}{2}i \end{bmatrix}$ y

$$\left(\frac{3}{2}-\frac{\sqrt{3}}{2}i\right)\begin{bmatrix} \frac{1}{2}+\frac{\sqrt{3}}{2}i \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{3}{2}+\frac{\sqrt{3}}{2}i \\ -\frac{3}{2}+\frac{\sqrt{3}}{2}i \end{bmatrix}. \checkmark$$

▶ **Ejemplo 5.54** Sea $A = \begin{bmatrix} i & -i & i \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$. Encontrar:

1. Los valores propios de A.
2. Los vectores propios correspondientes a cada valor propio.
3. Los espacios propios y sus respectivas dimensiones.◀

Solución 1. Valores propios:

$$\begin{vmatrix} i-\lambda & -i & i \\ 0 & 1-\lambda & 1 \\ 0 & -1 & 1-\lambda \end{vmatrix} = (i-\lambda)((1-\lambda)(1-\lambda)+1)$$
$$= (i-\lambda)(\lambda^2-2\lambda+2).$$

Por tanto,

$$\lambda = i \quad \text{y}$$
$$\lambda = \frac{2\pm\sqrt{4-8}}{2}$$
$$= \frac{2\pm\sqrt{4i^2}}{2}$$
$$= 1\pm i.$$

Así que los valores propios son $\lambda_1 = i$, $\lambda_2 = 1+i$ y $\lambda_3 = 1-i$.

2. Vectores propios. Para encontrar los vectores propios se tienen que encontrar las soluciones no triviales de los sistemas homogéneos $(A-\lambda_j I_3)\vec{u} = \vec{0}_{\mathbb{C}^3}$, para $j = 1,2,3$.

$\lambda_1 = i$:

$$\begin{bmatrix} 0 & -i & i \\ 0 & 1-i & 1 \\ 0 & -1 & 1-i \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & -1+i \\ 0 & 1-i & 1 \\ 0 & -i & i \end{bmatrix}$$
$$\sim \begin{bmatrix} 0 & 1 & -1+i \\ 0 & 0 & -2-2i \\ 0 & 0 & -1-i \end{bmatrix}$$
$$\sim \begin{bmatrix} 0 & 1 & -1+i \\ 0 & 0 & 1+i \\ 0 & 0 & 0 \end{bmatrix};$$

y por tanto

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} = r\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Así que los vectores propios correspondientes a λ_1 tienen la forma $(r,0,0)$ con $r \in \mathbb{C} - \{0\}$.

$\lambda_2 = 1+i$:

$$\begin{bmatrix} -1 & -i & i \\ 0 & -i & 1 \\ 0 & -1 & -i \end{bmatrix} \sim \begin{bmatrix} -1 & -i & i \\ 0 & -i & 1 \\ 0 & -1 & -i \end{bmatrix} \sim \begin{bmatrix} -1 & -i & i \\ 0 & -i & 1 \\ 0 & 0 & 0 \end{bmatrix};$$

por ende

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} (-1+i)r \\ -ir \\ r \end{bmatrix} = r \begin{bmatrix} (-1+i) \\ -i \\ 1 \end{bmatrix}.$$

Luego, los vectores propios correspondientes a λ_2 tienen la forma $r(-1+i,-i,1)$, con $r \in \mathbb{C}$ y $r \neq 0$.

$\lambda_3 = 1-i$:

$$\begin{bmatrix} -1+2i & -i & i \\ 0 & i & 1 \\ 0 & -1 & i \end{bmatrix} \sim \begin{bmatrix} -1+2i & -i & i \\ 0 & i & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

por lo que

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} (-1+3i)r \\ 5ir \\ 5r \end{bmatrix} = r \begin{bmatrix} (-1+3i) \\ 5i \\ 5 \end{bmatrix}.$$

Entonces los vectores propios correspondientes a λ_3 tienen la forma $r(-1+3i,5i,5)$ con $r \in \mathbb{C}$, $r \neq 0$.

3. Por lo precedente,

$$E_i = \text{gn}((1,0,0)),$$
$$E_{1+i} = \text{gn}((-1+i,-i,1)),$$
$$E_{1-i} = \text{gn}((-1+3i,5i,5));$$

y por ende,

$$\dim(E_i) = 1,$$
$$\dim(E_{1+i}) = 1,$$
$$\dim(E_{1-i}) = 1. \checkmark$$

Dado que un polinomio de grado n con coeficientes reales y/o complejos tiene n raíces complejas (recuerde que un número real se puede considerar como un número complejo), contando multiplicidades, se tiene la validez del siguiente teorema.

Teorema 5.35 *Sea A una matriz cuadrada de orden n (con componentes reales y/o complejas). Entonces A tiene exactamente n valores propios reales y/o complejos (contando multiplicidades, es decir raíces múltiples o repetidas del polinomio característico).*

► **Ejemplo 5.55**

1. Sea $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Entonces

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I_2) \\ &= \begin{vmatrix} 1-\lambda & 0 \\ 1 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)(1-\lambda) \\ &= 1-\lambda^2. \end{aligned}$$

Entonces los valores propios de A son $\lambda_1 = \lambda_2 = 1$, raíces repetidas ($\lambda = 1$ es raíz doble o de multiplicidad 2) del polinomio característico.

2. Sea

$$A = \begin{bmatrix} 3 & 0 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ -4 & 5 & i & 0 \\ 0 & 6 & -7 & i \end{bmatrix}.$$

Entonces

$$\det(A - \lambda I_4) = (3-\lambda)^2 (i-\lambda)^2.$$

Luego, los valores propios de A son $\lambda_1 = 3$, $\lambda_2 = 3$, $\lambda_3 = i$ y $\lambda_4 = i$. $\lambda = 3$ con multiplicidad 2 y $\lambda = i$ con multiplicidad 2.◄

Propiedades de los valores propios

Los valores propios tienen propiedades algebraicas directamente relacionadas con la matriz de la cual provienen.

Teorema 5.36 *Sea A una matriz cuadrada de orden n (con entradas reales y/o complejas) y sean* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *los valores propios de A (reales y/o complejos); es decir, las raíces del polinomio característico consideradas con distintos subíndices aun si algunas coinciden; es decir, si son raíces múltiples del polinomio característico. Sea*[30] $p_A(\lambda) = (-1)^n \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ *este polinomio. Entonces*

1. $$c_0 = \det(A) = \lambda_1 \lambda_2 \cdots \lambda_n.$$
2. $$\operatorname{tra}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

[30] Cfr. teorema 5.29, página 467.

DEMOSTRACIÓN ■ 1. Dado que λ_j, $j = 1, \ldots, n$, son las raíces del polinomio característico, se tiene

$$\begin{aligned} p_A(\lambda) &= (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) \\ &= (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda), \end{aligned} \tag{5.31}$$

y ya que

$$p_A(\lambda) = \det(A - \lambda I_n),$$

se sigue que

$$c_0 = \det(A) = p_A(0) = \lambda_1\lambda_2\cdots\lambda_n.$$

2. Del teorema 5.29 sabemos que $c_{n-1} = (-1)^{n-1}\text{tra}(A)$. Entonces

$$p_A(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}\text{tra}(A)\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \cdots + c_1\lambda + c_0.$$

Al desarrollar el lado derecho de (5.31) y comparar con la precedente igualdad se obtiene

$$(-1)^{n-1}\text{tra}(A) = (-1)^{n-1}(\lambda_1 + \lambda_2 + \cdots + \lambda_n);$$

de donde

$$\text{tra}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n. \quad ■$$

► **Ejemplo 5.56** Sea

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix},$$

la matriz del ejemplo 5.53. Ahí vimos que los valores propios de A son $\lambda_1 = \frac{3}{2} + \frac{\sqrt{3}}{2}i$ y $\lambda_2 = \frac{3}{2} - \frac{\sqrt{3}}{2}i$. Entonces, de acuerdo con el teorema 5.36,

$$\begin{aligned} \det(A) = \lambda_1\lambda_2 &= \left(\tfrac{3}{2} + \tfrac{\sqrt{3}}{2}i\right)\left(\tfrac{3}{2} - \tfrac{\sqrt{3}}{2}i\right) \\ &= \frac{9}{4} + \frac{3}{4} \\ &= 3 \end{aligned}$$

y

$$\begin{aligned} \text{tra}(A) &= \lambda_1 + \lambda_2 \\ &= \tfrac{3}{2} + \tfrac{\sqrt{3}}{2}i + \tfrac{3}{2} - \tfrac{\sqrt{3}}{2} \\ &= 3. \end{aligned}$$

Comprobación:

$$\det(A) = \begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} = 2 + 1 = 3,$$

$$\text{tra}(A) = a_{11} + a_{22} = 1 + 2 = 3. ◄$$

❍ **Nota 5.13**

1. Observe que si una matriz tiene componentes reales, su determinante y su traza son también números reales no importando si los valores propios son o no números complejos; sin embargo, el determinante es el producto de los valores propios y la traza es la suma de éstos. Entonces, aunque los valores propios de una matriz con componentes reales sean números complejos, su producto y suma son números reales como se ilustró en el ejemplo 5.56. ¡Sorprendente!, ¿verdad? En realidad no tanto, es un hecho conocido y fácil de probar, que las raíces complejas de polinomios con *coeficientes reales* aparecen en pares conjugados; i.e., si $z = a + bi$ es raíz, también $\bar{z} = a - bi$ es raíz. Entonces, ya que el producto y la suma de números complejos son asociativos y conmutativos, podemos arreglar $\lambda_1\lambda_2\cdots\lambda_n$ y $\lambda_1 + \lambda_2 + \cdots + \lambda_n$ de tal suerte que pares conjugados estén uno tras otro; de esta forma, puesto que $z\bar{z} = a^2 + b^2 \in \mathbb{R}$ y $z + \bar{z} = 2a \in \mathbb{R}$, $\lambda_1\lambda_2\cdots\lambda_n$ y $\lambda_1 + \lambda_2 + \cdots + \lambda_n$ serán números reales.
2. Por otro lado, si las componentes de la matriz son números complejos, entonces el determinante y la traza pueden o no ser números reales, y el polinomio característico tendrá posiblemente coeficientes complejos. Sin embargo, sigue siendo válido que la suma y el producto de los valores propios son, respectivamente, la traza y el determinante de dicha matriz. Veamos esto con algunos casos particulares: con la información dada en el ejemplo 5.53, $\text{tra}(A) = 2 + i = \lambda_1 + \lambda_2 + \lambda_3$, $\det(A) = 2i = \lambda_1\lambda_2\lambda_3$ y $p_A(\lambda) = -\lambda^3 + (2+i)\lambda^2 + (-2-2i)\lambda + 2i$, como el lector fácilmente puede comprobar. Mientras que con los datos del inciso 5.55 del ejemplo 5.55, $\text{tra}(A) = 6 + 2i$ y $\det(A) = -9$. Y, finalmente, si $A = \begin{bmatrix} 1 & i \\ 0 & 1 \end{bmatrix}$, $p_A(\lambda) = \lambda^2 - 2\lambda + 1$, $\det(A) = 1$, $\text{tra}(A) = 2$.

Diagonalización sobre el campo $\mathbb{C}$

Ya vimos que una matriz con componentes reales puede tener valores propios complejos; entonces no puede ser diagonalizable de acuerdo con la definición 5.22. Sin embargo, si se permite que los escalares y las componentes de los vectores propios sean números complejos en el sentido de la definición 5.24, es posible que este tipo de matrices sean "diagonalizables" en el campo $\mathbb{C}$; esto es, que existan un par de matrices C, D, con componentes complejas, tales que $C^{-1}AC$ sea una matriz diagonal.

► **Ejemplo 5.57** Si $A = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}$, entonces

$$\begin{aligned} p_A(\lambda) &= \begin{vmatrix} 1-\lambda & 2 \\ -1 & -1-\lambda \end{vmatrix} \\ &= -(1-\lambda^2) + 2 \\ &= \lambda^2 + 1. \end{aligned}$$

Así que los valores propios de A en $\mathbb{C}$ son $\lambda_1 = i$ y $\lambda_2 = -i$. Para $\lambda_1 = i$,

$$\begin{bmatrix} 1-i & 2 \\ -1 & -1-i \end{bmatrix} \sim \begin{bmatrix} 1 & 1+i \\ 1-i & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1+i \\ 0 & 0 \end{bmatrix};$$

por lo que los vectores propios para λ_1 tienen la forma

$$\alpha \begin{bmatrix} -1-i \\ 1 \end{bmatrix}, \quad \alpha \in \mathbb{C} - \{0\}.$$

Para $\lambda_2 = -i$,

$$\begin{bmatrix} 1+i & 2 \\ -1 & -1+i \end{bmatrix} \sim \begin{bmatrix} 1 & 1-i \\ 1+i & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1-i \\ 0 & 0 \end{bmatrix};$$

luego los valores propios correspondientes a λ_2 tienen la forma

$$\beta \begin{bmatrix} -1+i \\ 1 \end{bmatrix}, \quad \beta \in \mathbb{C} - \{0\}.$$

Entonces, si $C = \begin{bmatrix} -1-i & -1+i \\ 1 & 1 \end{bmatrix}$, se tiene $C^{-1} = \frac{1}{-2i} \begin{bmatrix} 1 & 1-i \\ -1 & -1-i \end{bmatrix}$ y

$$\begin{aligned} C^{-1}AC &= -\frac{1}{2i} \begin{bmatrix} 1 & 1-i \\ -1 & -1-i \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -1-i & -1+i \\ 1 & 1 \end{bmatrix} \\ &= -\frac{1}{2i} \begin{bmatrix} 1 & 1-i \\ -1 & -1-i \end{bmatrix} \begin{bmatrix} 1-i & 1+i \\ i & -i \end{bmatrix} \\ &= -\frac{1}{2i} \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{i} & 0 \\ 0 & \frac{1}{i} \end{bmatrix} \\ &= \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}. \end{aligned}$$

Es decir,

$$C^{-1}AC = \operatorname{diag}(i, -i). ◀$$

Es natural entonces extender el concepto de diagonalización sobre el campo $\mathbb{C}$.

Definición 5.25 *Sea $A \in \mathfrak{M}_n(\mathbb{C})$ una matriz cuadrada de orden n con entradas complejas. La matriz A es diagonalizable sobre el campo $\mathbb{C}$ si existe un par de matrices $C, D \in \mathfrak{M}_n(\mathbb{C})$, con C una matriz invertible y $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ una matriz diagonal, tales que*

$$D = C^{-1}AC.$$

Al par (C, D) se le llama entonces una diagonalización para la matriz A.

Toda la teoría relativa a la diagonalización que se desarrolló en el apartado 5.22 es válida para la diagonalización sobre el campo $\mathbb{C}$ y se transfiere directamente permitiendo que los escalares y las componentes de las matrices sean números complejos y que los vectores propios "habiten" en $\mathbb{C}^n$.

► **Ejemplo 5.58** Sea

$$A = \begin{bmatrix} i & -14+14i & 7+7i \\ 0 & 2-i & -1-i \\ 0 & -2-2i & -1+2i \end{bmatrix}.$$

Determinar si A es diagonalizable (en $\mathbb{C}$). En caso afirmativo encontrar una diagonalización para A.◀

Solución

$$\begin{aligned} p_A(\lambda) &= \begin{vmatrix} i-\lambda & -14+14i & 7+7i \\ 0 & 2-i-\lambda & -1-i \\ 0 & -2-2i & -1+2i-\lambda \end{vmatrix} \\ &= (i-\lambda)\begin{vmatrix} 2-i-\lambda & -1-i \\ -2-2i & -1+2i-\lambda \end{vmatrix} \\ &= (i-\lambda)\left((2-i-\lambda)(-1+2i-\lambda)-(-1-i)(-2-2i)\right) \\ &= (i-\lambda)\left(5i-\lambda-i\lambda+\lambda^2-4i\right) \\ &= (i-\lambda)(\lambda^2-(1+i)\lambda+i) \\ &= (i-\lambda)(\lambda-i)(\lambda-1) \\ &= (\lambda-i)^2(1-\lambda). \end{aligned}$$

Los valores propios distintos son entonces $\lambda_1 = i$, $\lambda_2 = 1$ con multiplicidades algebraicas $\mu_a(\lambda_1) = 2$ y $\mu_a(\lambda_1) = 1$. Resolvamos los sistemas $(A-\lambda_j I_3)\vec{u} = \vec{0}_{\mathbb{R}^3}$, $j = 1,2$.

Para λ_1:

$$\begin{bmatrix} 0 & -14+14i & 7+7i \\ 0 & 2-2i & -1-i \\ 0 & -2-2i & -1+i \end{bmatrix} \xleftrightarrow{R_2 \leftrightarrow -iR_1 - iR_2} \begin{bmatrix} 0 & -2+2i & 1+i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

por lo que $y = -(1+i)(-2+2i)^{-1}z = \frac{1}{2}iz$, y entonces

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \alpha \\ i\beta \\ 2\beta \end{bmatrix} = \alpha\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta\begin{bmatrix} 0 \\ i \\ 2 \end{bmatrix}; \alpha,\beta \in \mathbb{C}.$$

Para λ_2:

$$\begin{bmatrix} i-1 & -14+14i & 7+7i \\ 0 & 1-i & -1-i \\ 0 & -2-2i & -2+2i- \end{bmatrix} \xleftrightarrow{R_3 \leftrightarrow 2iR_2 + R_3} \begin{bmatrix} i-1 & -14+14i & 7+7i \\ 0 & 1-i & -1-i \\ 0 & 0 & 0 \end{bmatrix},$$

por lo que

$$y = (1+i)(1-i)^{-1}z = iz$$

y

$$\begin{aligned} x &= ((14-14i)y + (-7-7i)z)(i-1)^{-1} \\ &= ((14-14i)iz + (-7-7i)z)(i-1)^{-1} \\ &= -7iz; \end{aligned}$$

es decir,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -7i\gamma \\ i\gamma \\ \gamma \end{bmatrix} = \gamma\begin{bmatrix} -7i \\ i \\ 1 \end{bmatrix}; \gamma \in \mathbb{C}.$$

Entonces

$$\begin{aligned} \mu_g(\lambda_1) &= \dim(E_i) = 2, \\ \mu_g(\lambda_2) &= \dim(E_1) = 1. \end{aligned}$$

Ya que $\mu_g(\lambda_1) = 2 = \mu_a(\lambda_1)$ y $\mu_g(\lambda_2) = 1 = \mu_a(\lambda_2)$, A es diagonalizable y una diagonalización para esta matriz es

$$C = \begin{bmatrix} 1 & 0 & -7i \\ 0 & i & i \\ 0 & 2 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Comprobación:

$$\begin{aligned} AC &= \begin{bmatrix} i & -14+14i & 7+7i \\ 0 & 2-i & -1-i \\ 0 & -2-2i & -1+2i \end{bmatrix} \begin{bmatrix} i & 0 & -7i \\ 0 & -1 & i \\ 0 & 2i & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 & -7i \\ 0 & -i & i \\ 0 & -2 & 1 \end{bmatrix} \end{aligned}$$

y

$$\begin{aligned} CD &= \begin{bmatrix} i & 0 & -7i \\ 0 & -1 & i \\ 0 & 2i & 1 \end{bmatrix} \begin{bmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 & -7i \\ 0 & -i & i \\ 0 & -2 & 1 \end{bmatrix}, \end{aligned}$$

lo cual implica que $AC = CD$; es decir, $C^{-1}AC = D$. ✔

▶ **Ejemplo 5.59** Sea $A = \begin{bmatrix} i & 1 \\ 0 & i \end{bmatrix}$. Determinar si A es diagonalizable.◀

Solución

$$p_A(\lambda) = \begin{vmatrix} i-\lambda & 1 \\ 0 & i-\lambda \end{vmatrix} = (i-\lambda)^2.$$

Entonces los valores propios de A son $\lambda = \lambda_1 = \lambda_2 = i$ con multiplicidad algebraica $\mu_a(\lambda) = 2$. Puesto que

$$A - \lambda I_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

los vectores propios de λ están dados por

$$\begin{bmatrix} x \\ y \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \alpha \in \mathbb{C} - \{0\}.$$

Por lo que $\dim(E_i) = 1$; y por tanto, $\mu_a(\lambda) \neq \mu_g(\lambda)$, así que A no es diagonalizable. ✔

5.3.4 *Operadores autoadjuntos y matrices simétricas*

En la práctica difícilmente podremos evitar encontrar matrices reales que tengan valores propios complejos; sin embargo, existe un tipo especial de matrices para las cuales todos los valores propios son

números reales. Éstas son las matrices simétricas, que aparecen con mucha frecuencia en aplicaciones. El objetivo de este último apartado es probar que efectivamente los valores propios de cualquier matriz simétrica en $\mathbb{R}$ son números reales y, más aún, que toda matriz simétrica con componentes reales es diagonalizable. Para alcanzar este fin, de la manera más breve y directa, necesitamos trabajar en espacios complejos y extender el concepto de producto interior a este tipo de espacios. Invitamos nuevamente al lector a consultar los apartados 1.1.5, 1.2.8, 2.1.5, 2.2.5, 3.5 y la sección B.1 del apéndice B para la mejor comprensión de este segmento.

Definición 5.26 *Sea* **E** *un espacio vectorial sobre el campo de los números complejos. Un* ***producto interior*** *(o producto escalar) en* **E** *es una función que a cada par de vectores* $\vec{u},\vec{v}\in\mathbf{E}$ *les asigna un número complejo denotado por* $\langle\vec{u},\vec{v}\rangle$*, tal que:*

1. $\langle\vec{u},\vec{v}\rangle=\overline{\langle\vec{v},\vec{u}\rangle},\quad\forall\vec{u},\vec{v}\in\mathbf{E}.$
2. $\langle\vec{u},\vec{v}+\vec{w}\rangle=\langle\vec{u},\vec{v}\rangle+\langle\vec{u},\vec{w}\rangle,\quad\forall\vec{u},\vec{v},\vec{w}\in\mathbf{E}.$
3. $\langle\alpha\vec{u},\vec{v}\rangle=\alpha\langle\vec{u},\vec{v}\rangle,\quad\forall\vec{u},\vec{v}\in\mathbf{E},\ \forall\alpha\in\mathbb{C}.$
4. $\langle\vec{u},\vec{u}\rangle\geq 0,\quad\forall\vec{u}\in\mathbf{E}.$
5. $\langle\vec{u},\vec{u}\rangle=0\Leftrightarrow\vec{u}=\vec{0}_{\mathbf{E}}.$

Observe que la definición 5.26 es exactamente la misma que la definición 4.1 (cfr. pág. 236) de producto interior en espacios reales salvo la propiedad 1 que difiere de la propiedad de simetría en espacios reales; pues en la propiedad 1 del producto interior en espacios complejos está involucrado el conjugado de un número complejo.[31] Por esta razón es que la definición precedente tiene como caso particular a la definición 4.1 al restringirse a espacios reales; pues $z=\overline{z}$ si $z\in\mathbb{R}$.

❍ **Nota 5.14**

1. Si $\vec{u}=(z_1,\ldots,z_n)\in\mathbb{C}^n$, utilizaremos la notación $\overline{\vec{u}}$ para representar el vector de $\mathbb{C}^n$ que tiene por componentes los conjugados de las componentes del vector $\vec{u}$; i.e., $\overline{\vec{u}}=(\overline{z_1},\ldots,\overline{z_n})$. De manera análoga, si $A\in\mathfrak{M}_{m\times n}(\mathbb{C})$ es una matriz, representamos a la matriz cuyas componentes son los conjugados de la matriz A por el símbolo $\overline{A}$.
2. Es fácil mostrar que si $A\in\mathfrak{M}_{m\times n}(\mathbb{C})$ y $\vec{u}\in\mathbb{C}^n$, entonces $\overline{A\vec{u}}=\overline{A}\,\overline{\vec{u}}$. Lo cual se deja de ejercicio al lector.
3. Si A es una matriz con componentes reales, es evidente que $\overline{A}=A$.

► **Ejemplo 5.60** No es difícil probar que si se define para cada $\vec{u}=(z_1,\ldots,z_n)\in\mathbb{C}^n$ y $\vec{v}=(w_1,\ldots,w_n)\in\mathbb{C}^n$,

$$\langle\vec{u},\vec{v}\rangle=\vec{u}\cdot\vec{v}=\sum_{j=1}^{n}z_j\overline{w_j},$$

entonces $\langle\vec{u},\vec{v}\rangle$ es un producto interior en $\mathbb{C}^n$. Se le llama el **producto interior canónico en** $\mathbb{C}^n$. Utilizando la notación que convenimos en 5.14, este producto interior se puede escribir como el producto matricial

$$\langle\vec{u},\vec{v}\rangle=\vec{u}^t\,\overline{\vec{v}}.$$ ◄

[31] Recuerde que si $z=a+bi\in\mathbb{C}$, el conjugado de z es $\overline{z}=a-bi$.

▶ **Ejemplo 5.61**

$$\begin{aligned}\langle(1-i,1,i),(2+i,1+i,2)\rangle &= \begin{bmatrix}1-i & 1 & i\end{bmatrix}\overline{\begin{bmatrix}2+i\\1\\i\end{bmatrix}}\\ &= \begin{bmatrix}1-i & 1 & i\end{bmatrix}\begin{bmatrix}2-i\\1\\-i\end{bmatrix}\\ &= (1-i)(2-i)+(1)(1)+(i)(-i)\\ &= 3-3i.\end{aligned}$$ ◀

▶ **Ejemplo 5.62** En $\mathfrak{M}_{m\times n}(\mathbb{C})$ se define, para cada $A,B\in\mathfrak{M}_{m\times n}(\mathbb{C})$,

$$\langle A,B\rangle = \operatorname{tra}(\overline{B}^t A).$$

Se puede probar, de manera análoga a como se hizo en el caso real (cfr. ejemplo 4.8 pág. 242), que éste es un producto interior en $\mathfrak{M}_{m\times n}(\mathbb{C})$.◀

Hay una pequeña, pero notable característica del producto interior en espacios complejos que es importante que el lector tenga presente; la primera propiedad del siguiente teorema, cuya demostración es sencilla, se deja como ejercicio al lector.

Teorema 5.37 *Sea* $\mathbf{E}$ *un espacio complejo y* $\langle\cdot,\cdot\rangle$ *un producto interior en este espacio. Entonces*

1. $\langle\vec{u},\alpha\vec{v}\rangle = \overline{\alpha}\,\langle\vec{u},\vec{v}\rangle$ *para todo* $\alpha\in\mathbb{C}$ *y para todo par de vectores* $\vec{u},\vec{v}\in\mathbf{E}$.
2. $\langle\vec{u}+\vec{v},\vec{w}\rangle = \langle\vec{u}+\vec{w},\vec{v}+\vec{w}\rangle\ \forall\vec{u},\vec{v},\vec{w}\in\mathbf{E}$.

Definición 5.27 *Si* $\mathbf{E}$ *es un espacio vectorial complejo con producto interior* $\langle\cdot,\cdot\rangle$, *se define, para cada* $\vec{u}\in\mathbf{E}$,

$$\|\vec{u}\| = \langle\vec{u},\vec{u}\rangle^{1/2}$$

y se le llama la ***norma*** *del vector* $\vec{u}$ ***inducida*** *por el producto interior.*

Se puede probar, de manera análoga al caso real, que el producto interior en un espacio vectorial complejo satisface la desigualdad de Schwarz, $|\langle\vec{u},\vec{v}\rangle|\le\|\vec{u}\|\,\|\vec{v}\|\ \forall\vec{u},\vec{v}\in\mathbf{E}$, y que la norma inducida es en efecto una norma en el sentido de la definición 4.11. Se define ortogonalidad de manera análoga al caso real; esto es, **dos vectores son ortogonales si su producto interior es nulo**.

Para abreviar diremos que un espacio vectorial real o complejo con producto interior $\langle\cdot,\cdot\rangle$ definido en él es un espacio **euclidiano**. El teorema de ortogonalización de Gram-Schmidt es también válido en espacios euclidianos complejos.

Si $\mathbf{E}$ es un espacio complejo, como en el caso real, un operador lineal en él es una transformación $T:\mathbf{E}\to\mathbf{E}$ que satisface:

- $T(\vec{u}+v) = T(\vec{u})+T(\vec{v})\ \forall\vec{u},\vec{v}\in\mathbf{E}$.
- $T(\alpha\vec{u}) = \alpha T(\vec{u})\ \forall\vec{u}\in\mathbf{E},\forall\alpha\in\mathbb{C}$.

Los operadores lineales en espacios complejos tienen las mismas propiedades que los operadores lineales en espacios reales, pero trabajando con escalares complejos; así que haremos uso libre de éstas en todo lo que sigue.

En el caso de espacios euclidianos, los valores propios y vectores propios correspondientes de un operador lineal están relacionados con el producto interior, como hacemos patente en la siguiente proposición.

Teorema 5.38 *Sean* $\mathbf{E}$ *un espacio euclidiano y* $T:\mathbf{E}\to\mathbf{E}$ *un operador lineal. Si* λ *es un valor propio de* T *con vector propio correspondiente* $\vec{u}$*, entonces*

$$\lambda=\frac{\langle T(\vec{u}),\vec{u}\rangle}{\|\vec{u}\|^2}. \tag{5.32}$$

DEMOSTRACIÓN ■

$$\begin{aligned}\langle T(\vec{u}),\vec{u}\rangle &= \langle\lambda\vec{u},\vec{u}\rangle\\ &= \lambda\langle\vec{u},\vec{u}\rangle,\end{aligned}$$

de donde se tiene (5.32). ■

Definición 5.28 *Sea* $\mathbf{E}$ *un espacio euclidiano. Se dice que un operador lineal* $T:\mathbf{E}\to\mathbf{E}$ *es* ***autoadjunto*** *(o hermitiano) si*

$$\langle T(\vec{u}),\vec{v}\rangle=\langle\vec{u},T(\vec{v})\rangle\quad\forall\vec{u},\vec{v}\in\mathbf{E}.$$

► **Ejemplo 5.63** Sea $A\in\mathfrak{M}_n$ una matriz *simétrica* con componentes reales, entonces T_A, con $T_A(\vec{u})=A\vec{u}\ \forall\vec{u}\in\mathbb{C}^n$, es un operador autoadjunto en $\mathbb{C}^n$ con el producto interior canónico en este espacio (cfr. ejemplo 5.60). En efecto, como A es simétrica, $A^t=A$; y como A es real , $\overline{A}=A$; luego

$$\begin{aligned}\langle T_A(\vec{u}),\vec{v}\rangle &= (T_A(\vec{u}))^t\,\overline{\vec{v}}\\ &= (A\vec{u})^t\overline{\vec{v}}\\ &= (\vec{u})^tA^t\overline{\vec{v}}\\ &= (\vec{u})^tA\overline{\vec{v}}\\ &= \vec{u}^t\overline{A}\,\overline{\vec{v}}\\ &= \vec{u}^t\overline{A\vec{v}}\\ &= \langle\vec{u},A\vec{v}\rangle\\ &= \langle\vec{u},T_A(\vec{v})\rangle.\end{aligned}$$ ◄

Mostremos a continuación que los valores propios de un operador autoadjunto y, en particular, de una matriz real simétrica, son todos números reales.

Teorema 5.39

1. *Sean* $\mathbf{E}$ *un espacio euclidiano y* $T:\mathbf{E}\to\mathbf{E}$ *un operador lineal autoadjunto. Si* λ *es un valor propio de* T*, entonces* λ *es real.*
2. *Sea* A *una matriz simétrica cuyas componentes son números reales, entonces los valores propios de* A *son (todos) números reales.*

DEMOSTRACIÓN ■ 1. Sea λ un valor propio de T con vector propio correspondiente $\vec{u} \in \mathbf{E}$, entonces, por (5.32) del teorema 5.38,

$$\begin{aligned}\lambda &= \frac{\langle T(\vec{u}), \vec{u}\rangle}{\|\vec{u}\|^2} \\ &= \frac{\langle \vec{u}, T(\vec{u})\rangle}{\|\vec{u}\|^2} \\ &= \frac{\overline{\langle T(\vec{u}), \vec{u}\rangle}}{\|\vec{u}\|^2} \\ &= \overline{\lambda}.\end{aligned}$$

Puesto que $\lambda = \overline{\lambda}$ se concluye[32] que $\lambda \in \mathbb{R}$.

2. Por el ejemplo 5.63, el operador T_A es autoadjunto y tiene los mismos valores propios que A; así que esta afirmación es consecuencia del inciso precedente. (Recuerde que una matriz cuadrada de orden n tiene n valores propios en $\mathbb{C}$ contando multiplicidades.) ■

▶ **Ejemplo 5.64** Sea $A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & -1 & 1 \\ 2 & 1 & 2 \end{bmatrix}$. El lector puede verificar, efectuando el cálculo, que

$$p_A(\lambda) = \lambda^3 - 2\lambda^2 - 7\lambda + 5.$$

Haciendo división sintética o evaluando es fácil comprobar que ± 1 y ± 5 no son raíces enteras de este polinomio; así el polinomio característico no tiene raíces racionales. Sin embargo, puesto que A es simétrica, todos sus valores propios deben ser reales, como se hace patente en la figura 5-5.◀

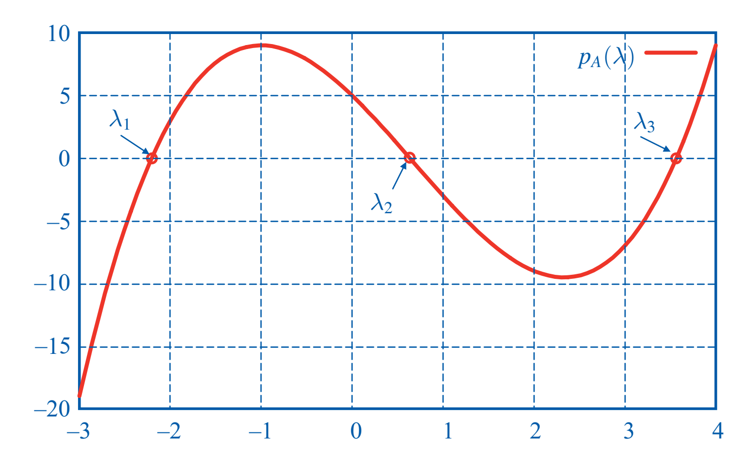


Figura 5-5 • Gráfica del polinomio característico $p_A(\lambda)$ que interseca al eje x en tres puntos: los valores característicos de la matriz A. Aunque no es simple calcular estos valores en forma exacta, existen y son números reales por ser la matriz simétrica.

Una propiedad trascendente de los operadores autoadjuntos es que los vectores propios que corresponden a valores propios distintos de este tipo de operadores son ortogonales. Probamos a continuación este importante resultado.

[32] $z = a + bi = \overline{z} = a - bi \Rightarrow 2bi = 0 \Rightarrow b = 0 \Rightarrow z = a \Rightarrow z \in \mathbb{R}$.

Teorema 5.40 *Sean* $\mathbf{E}$ *un espacio euclidiano y* $T:\mathbf{E}\to\mathbf{E}$ *un operador lineal autoadjunto. Si* λ_1 *y* λ_2 *son valores propios diferentes de* T *con vectores propios correspondientes* $\vec{u}$ *y* $\vec{v}$*, respectivamente, entonces* $\vec{u}$ *y* $\vec{v}$ *son ortogonales; esto es,* $\langle\vec{u},\vec{v}\rangle = 0$.

DEMOSTRACIÓN ■ Como T es autoadjunto, por el inciso 5.39 del teorema 5.39, $\lambda_1,\lambda_2\in\mathbb{R}$. Entonces

$$\begin{aligned}\langle\lambda_1\vec{u},\vec{v}\rangle &= \langle T(\vec{u}),\vec{v}\rangle\\ &= \langle\vec{u},T(\vec{v})\rangle\\ &= \langle\vec{u},\lambda_2\vec{v}\rangle\\ &= \overline{\lambda_2}\,\langle\vec{u},\vec{v}\rangle\\ &= \lambda_2\,\langle\vec{u},\vec{v}\rangle,\end{aligned}$$

de donde

$$\lambda_1\,\langle\vec{u},\vec{v}\rangle = \lambda_2\,\langle\vec{u},\vec{v}\rangle;$$

y por tanto,

$$(\lambda_1-\lambda_2)\,\langle\vec{u},\vec{v}\rangle = 0;$$

y ya que $\lambda_1\neq\lambda_2$, se tiene

$$\langle\vec{u},\vec{v}\rangle = 0. \quad ■$$

Mostramos a continuación el resultado más importante de esta sección y que tiene como caso particular el hecho de que toda matriz real simétrica es diagonalizable.

Teorema 5.41 *Sean* $\mathbf{E}$ *un espacio euclidiano de dimensión finita y* $T:\mathbf{E}\to\mathbf{E}$ *un operador lineal autoadjunto. Entonces existe una base ortonormal del espacio* $\mathbf{E}$ *formada por vectores propios de* T*. En particular, en virtud del teorema 5.25,* T *es diagonalizable.*

DEMOSTRACIÓN ■ Procederemos por inducción sobre la dimensión del espacio $\mathbf{E}$. Si $n=1$, sea $\vec{u}\in\mathbf{E}$ con $\|\vec{u}\|=1$; entonces $\mathscr{B}=\{\vec{u}\}$ es una base ortonormal de $\mathbf{E}$ y $T(\vec{u})=\lambda\vec{u}$ para algún escalar $\lambda\in\mathbb{R}$; por lo que $\vec{u}$ es un vector propio de T. Supongamos que el resultado es cierto para cualquier espacio vectorial de dimensión $n-1$. Sea λ_1 un valor propio de T (λ_1 existe porque $\mathbf{E}$ tiene dimensión finita y es real porque T es autoadjunto) y sea $\vec{u}_1\in\mathbf{E}$ un vector propio correspondiente con $\|\vec{u}_1\|=1$. Sean $S=\text{gn}(\vec{u}_1)$ y

$$S^{\perp} = \{\vec{v}\in\mathbf{E}\,|\,\langle\vec{v},\vec{u}_1\rangle = 0\}.$$

Es decir, $S^{\perp}$ es el complemento ortogonal de S. Es fácil probar que $S^{\perp}$ es un subespacio[33] de $\mathbf{E}$. Completemos[34] la base $\mathscr{B}$ a una base ortonormal $\{\vec{u}_1,\vec{v}_2,\ldots,\vec{v}_n\}$ de $\mathbf{E}$ (si es necesario se debe aplicar el

[33] Cfr. ejercicio propuesto 95 del capítulo 4.

[34] Cfr. el procedimiento dado para este fin en la página 167.

proceso de ortogonalización de Gram-Schmidt en esta construcción). Sean $\vec{v} \in S^{\perp}$ y $\alpha_1, \alpha_2, \ldots, \alpha_n$ escalares tales que

$$\vec{v} = \alpha_1 \vec{u}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n,$$

entonces

$$0 = \langle \vec{v}, \vec{u}_1 \rangle = \alpha_1 \langle \vec{u}_1, \vec{u}_1 \rangle + \alpha_2 \langle \vec{v}_2, \vec{u}_1 \rangle + \cdots + \alpha_n \langle \vec{v}_n, \vec{u}_1 \rangle = \alpha_1,$$

pues $\{\vec{u}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ es ortonormal. Entonces

$$\vec{v} = \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n;$$

y puesto que los vectores $\vec{v}_j \in S^{\perp}$ ($\{\vec{u}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ son un conjunto ortonormal y por tanto, $(\vec{v}_i \perp \vec{u}_1\ \forall j)$, se concluye que

$$S^{\perp} = \mathrm{gn}(\vec{v}_2, \ldots, \vec{v}_n);$$

y ya que los vectores $\vec{v}_j$ son L.I., se tiene

$$\dim(S^{\perp}) = n - 1.$$

Si $\vec{v} \in S^{\perp}$, entonces $\langle \vec{v}, \vec{u}_1 \rangle = 0$; y dado que T es autoadjunto,

$$\langle \vec{u}_1, T(\vec{v}) \rangle = \langle T(\vec{u}_1), \vec{v} \rangle = \langle \lambda_1 \vec{u}_1, \vec{v} \rangle = \lambda_1 \langle \vec{u}_1, \vec{v} \rangle = 0,$$

así que $T(\vec{v}) \in S^{\perp}$; por lo que T es un operador lineal autoadjunto en el espacio vectorial $S^{\perp}$ de dimensión $n-1$. Por hipótesis de inducción existe una base ortonormal de vectores propios de T para este espacio, digamos $\{\vec{u}_2, \ldots, \vec{u}_n\}$. Entonces

$$\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$$

es una base ortonormal de vectores propios para $\mathbf{E}$. ■

Como corolario a este teorema tenemos el resultado final de este capítulo.

Teorema 5.42 *Si A es una matriz simétrica real, entonces A es diagonalizable.*

DEMOSTRACIÓN ■ Puesto que A es una matriz simétrica con componentes reales, entonces, por el ejemplo 5.63, el operador lineal T_A es autoadjunto y la afirmación es consecuencia inmediata del teorema 5.41. ■

5.4 Ejercicios resueltos y ejercicios propuestos

5.4.1 Ejercicios resueltos

Transformaciones lineales

1 Sean $\mathbf{E}$ un espacio vectorial de dimensión finita n y $\mathbf{F}$ un espacio vectorial arbitrario. Probar que si $\{\vec{e}_1, \ldots, \vec{e}_n\}$ es una base de $\mathbf{E}$ y $\vec{f}_1, \ldots, \vec{f}_n$ son vectores cualesquiera en $\mathbf{F}$, entonces existe una única transformación lineal $T : \mathbf{E} \to \mathbf{F}$ tal que $T(\vec{e}_i) = \vec{f}_i$ para cada $i = 1, \ldots, n$.

DEMOSTRACIÓN ■ Sea $\vec{u} \in \mathbf{E}$ cualquier vector; dado que $\{\vec{e}_1, \ldots, \vec{e}_n\}$ es una base de $\mathbf{E}$, existen escalares únicos α_i tales que $\vec{u} = \sum_{i=1}^{n} \alpha_i \vec{e}_i$. Se define $T : \mathbf{E} \to \mathbf{F}$ como

$$T(\vec{u}) = \sum_{i=1}^{n} \alpha_i \vec{f}_i.$$

Por la unicidad de los escalares se tiene que la transformación T está bien definida. Entonces

(a) Claramente $T(\vec{e}_i) = \vec{f}_i$ para todo $i = 1, \ldots, n$. $(\vec{e}_i = 0\vec{e}_1 + \cdots + 0\vec{e}_{i-1} + 1\vec{e}_i + 0\vec{e}_{i+1} + \cdots + 0\vec{e}_n)$

(b) Si $a, b \in \mathbb{R}$ y $\vec{u} = \sum_{i=1}^{n} \alpha_i \vec{e}_i$, $\vec{v} = \sum_{i=1}^{n} \beta_i \vec{e}_i$ son un par de vectores en $\mathbf{E}$,

$$\begin{aligned}
T(a\vec{u} + b\vec{v}) &= T\left(a\sum_{i=1}^{n} \alpha_i \vec{e}_i + b\sum_{i=1}^{n} \beta_i \vec{e}_i\right) \\
&= T\left(\sum_{i=1}^{n} a\alpha_i \vec{e}_i + \sum_{i=1}^{n} b\beta_i \vec{e}_i\right) \\
&= T\left(\sum_{i=1}^{n} (a\alpha_i \vec{e}_i + b\beta_i \vec{e}_i)\right) \\
&= T\left(\sum_{i=1}^{n} (a\alpha_i + b\beta_i)\, \vec{e}_i\right) \\
&= \sum_{i=1}^{n} (a\alpha_i + b\beta_i)\, \vec{f}_i \\
&= a\sum_{i=1}^{n} \alpha_i \vec{f}_i + b\sum_{i=1}^{n} \beta_i \vec{f}_i \\
&= aT(\vec{u}) + bT(\vec{v}).
\end{aligned}$$

Por tanto, $T \in \mathscr{L}(\mathbf{E}, \mathbf{F})$.

(c) Sea $T_1 \in \mathscr{L}(\mathbf{E}, \mathbf{F})$ tal que $T_1(\vec{e}_i) = \vec{f}_i$ para todo $i = 1, \ldots, n$. Si $\vec{u} = \sum_{i=1}^{n} \alpha_i \vec{e}_i$ es cualquier vector de $\mathbf{E}$, entonces

$$\begin{aligned}
T_1(\vec{u}) &= T_1\left(\sum_{i=1}^{n} \alpha_i \vec{e}_i\right) \\
&= \sum_{i=1}^{n} \alpha_i T_1(\vec{e}_i) \\
&= \sum_{i=1}^{n} \alpha_i \vec{f}_i \\
&= T(\vec{u});
\end{aligned}$$

y por tanto, $T = T_1$. ■

2 Encontrar la transformación lineal $T : \mathbf{P}_2 \to \mathbf{P}_1$ tal que $T(1) = 0$, $T(x) = 1$ y $T(x^2) = 2x$.

Solución El conjunto $\{1, x, x^2\}$ es una base de $\mathbf{P}_2$, por el ejercicio precedente $T : \mathbf{P}_2 \to \mathbf{P}_1$ definida, para cada $a + bx + cx^2 \in \mathbf{P}_2$, por

$$\begin{aligned} T(a+bx+cx^2) &= a\cdot 0+b\cdot 1+c\cdot 2x \\ &= b+2cx \end{aligned}$$

es lineal. Note que $T(p)=p'$. ✔

3 (**Transformación rotación**). Encontrar la transformación lineal $T:\mathbb{R}^2\to\mathbb{R}^2$ que a cada vector $\vec{u}$ lo transforme en el vector que se obtiene girando 45° a $\vec{u}$ en sentido contrario a las manecillas del reloj y que tiene la misma norma de $\vec{u}$.

Solución Por el ejercicio 1, $T:\mathbb{R}^2\to\mathbb{R}^2$, con $T(1,0)=(\sqrt{2}/2,\sqrt{2}/2)$, $T(0,1)=(-\sqrt{2}/2,\sqrt{2}/2)$ y si $(x,y)\in\mathbb{R}^2$,

$$\begin{aligned} T(x,y) &= x(\sqrt{2}/2,\sqrt{2}/2)+y(-\sqrt{2}/2,\sqrt{2}/2) \\ &= \frac{\sqrt{2}}{2}(x-y,x+y) \end{aligned}$$

es una transformación lineal. Sea $\vec{w}=T(x,y)=\frac{\sqrt{2}}{2}(x-y,x+y)$, entonces

$$\begin{aligned} \|\vec{w}\| &= \|\frac{\sqrt{2}}{2}(x-y,x+y)\| \\ &= \frac{\sqrt{2}}{2}\sqrt{2(x^2+y^2)} \\ &= \sqrt{x^2+y^2} \\ &= \|\vec{u}\|, \end{aligned}$$

$$\begin{aligned} \vec{u}\cdot\vec{w} &= (x,y)\cdot\frac{\sqrt{2}}{2}(x-y,x+y) \\ &= \frac{\sqrt{2}}{2}(x,y)\cdot(x-y,x+y) \\ &= \frac{\sqrt{2}}{2}(x^2+y^2); \end{aligned}$$

y por tanto

$$\begin{aligned} \frac{\vec{u}\cdot\vec{w}}{\|\vec{u}\|\,\|\vec{v}\|} &= \frac{\frac{\sqrt{2}}{2}(x^2+y^2)}{x^2+y^2} \\ &= \frac{\sqrt{2}}{2}. \end{aligned}$$

Así que el ángulo entre $\vec{u}$ y su imagen $\vec{w}=T(\vec{u})$ es

$$\phi=\arccos\left(\frac{\sqrt{2}}{2}\right)=45^\circ. \quad ✔$$

4 Sean $\mathbf{E}$ y $\mathbf{F}$ espacios vectoriales, con $\dim(\mathbf{E})=n$, y $\vec{f}_1,\ldots,\vec{f}_m$ vectores dados de $\mathbf{F}$. Probar que existe una transformación lineal $T:\mathbf{E}\to\mathbf{F}$ cuya imagen está generada por los vectores $\vec{f}_1,\ldots,\vec{f}_m$.

DEMOSTRACIÓN ■ Sea $\{\vec{e}_1, \ldots, \vec{e}_n\}$ una base de $\mathbf{E}$.

(a) Si $n \le m$, por el ejercicio 1 existe una transformación lineal $T : \mathbf{E} \to \mathbf{F}$ tal que $T(\vec{e}_i) = \vec{f}_i$ para cada $i = 1, \ldots, n$. Por el teorema 5.7 (cfr. pág. 429), dado que los vectores $\vec{e}_i$ generan a $\mathbf{E}$, los vectores $T(\vec{e}_i) = \vec{f}_i$, $i = 1, \ldots, n$, generan la imagen de T; por tanto, también los vectores $\vec{f}_i$, $i = 1, \ldots, m$.

(b) Si $n > m$, sean $\vec{f}_j = \vec{0}_{\mathbf{F}}$ para $j = m+1, \ldots, n$; por el ejercicio 1 existe una transformación lineal $T : \mathbf{E} \to \mathbf{F}$ tal que $T(\vec{e}_i) = \vec{f}_i$ para $i = 1, \ldots, n$. Puesto que los vectores $\vec{e}_i$ generan a $\mathbf{E}$, por el teorema 5.7, los vectores $\vec{f}_i$, $i = 1, \ldots, n$, generan a la imagen de T y, por tanto, también los vectores $\vec{f}_i$, $i = 1, \ldots, m$. ■

5 Encontrar una transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que su imagen esté generada por los vectores $(-1, 2, 1)$ y $(1, 1, 3)$.

Solución Sean $T(1,0,0) = (-1,2,1)$, $T(0,1,0) = (1,1,3)$ y $T(0,0,1) = (0,0,0)$. Se define entonces

$$\begin{aligned} T(x,y,z) &= xT(1,0,0) + yT(0,1,0) + zT(0,0,1) \\ &= x(-1,2,1) + y(1,1,3) + z(0,0,0) \\ &= (-x+y, 2x+y, x+3y) \end{aligned}$$

Por el ejercicio T es lineal y su imagen está generada por los vectores $(-1,2,1)$ y $(1,1,3)$. ✔

6 Sean $\mathbf{E}$ y $\mathbf{F}$ espacios vectoriales. Si la dimensión de $\mathbf{E}$ es infinita y $T \in \mathscr{L}(\mathbf{E}, \mathbf{F})$, probar que por lo menos uno de los subespacios $\operatorname{Ker}(T)$ o $T(\mathbf{E})$ tiene también dimensión infinita.

DEMOSTRACIÓN ■ Suponga que ambos subespacios tienen dimensión finita. Sean $\{\vec{e}_1, \ldots, \vec{e}_r\}$ y $\{T(\vec{f}_1), \ldots, T(\vec{f}_m)\}$ bases de $\operatorname{Ker}(T)$ y $T(E)$, respectivamente. Sea $\vec{u} \in \mathbf{E}$ cualquier vector, entonces existen escalares β_i tales que $T(\vec{u}) = \sum_{i=1}^m \beta_i T(\vec{f}_i)$, y por tanto, $T(\vec{u} - \sum_{i=1}^m \beta_i \vec{f}_i) = \vec{0}_{\mathbf{F}}$; así que $\vec{u} - \sum_{i=1}^m \beta_i \vec{f}_i \in \operatorname{Ker}(T)$; luego existen escalares α_i tales que $\vec{u} - \sum_{i=1}^m \beta_i \vec{f}_i = \sum_{i=1}^r \alpha_i \vec{e}_i$ y por ende,

$$\vec{u} = \sum_{i=1}^r \alpha_i \vec{e}_i + \sum_{i=1}^m \beta_i \vec{f}_i.$$

De donde se concluye que $\mathbf{E}$ está generado por los vectores $\vec{e}_1, \ldots, \vec{e}_r$, $\vec{f}_1, \ldots, \vec{f}_m$; lo cual es una contradicción a la hipótesis de que $\mathbf{E}$ tiene dimensión infinita. Por tanto, uno de los dos subespacios, al menos, debe tener dimensión infinita. ■

7 Sea $f \in C[a,b]$ y

$$g(x) = \int_a^b f(t)\cos(x-t)\,dt$$

para cada $x \in [a,b]$.

(a) Mostrar que $g \in C[a,b]$.

(b) Sea $T : C[a,b] \to C[a,b]$ definido, para cada $f \in C[a,b]$, por la función $T(f)$ donde

$$T(f)(x) = \int_a^b f(t)\cos(x-t)\,dt$$

para cada $x \in [a,b]$. Demostrar que T es lineal.

(c) Encontrar una base y la dimensión de la imagen de T.

(d) Determinar la dimensión del núcleo de T.

Solución *(a)* Para todo $x \in [a,b]$,

$$\begin{aligned} g(x) &= \int_a^b f(t)\cos(x-t)dt \\ &= \int_a^b f(t)(\cos x \cos t + \operatorname{sen} x \operatorname{sen} t)dt \quad \text{es decir,} \end{aligned}$$

$$g(x) = \left(\int_a^b f(t)\cos t\,dt\right)\cos x + \left(\int_a^b f(t)\operatorname{sen} t\,dt\right)\operatorname{sen} x \tag{5.33}$$

y, por tanto, dado que las funciones $y = \cos x$ y $y = \operatorname{sen} x$ son continuas en todo punto, se tiene que $g \in C[a,b]$.

(b) Sean $f_1, f_2 \in C[a,b]$ y $\alpha, \beta \in \mathbb{R}$, entonces, para todo $x \in C[a,b]$,

$$\begin{aligned} T(\alpha f_1 + \beta f_2)(x) &= \int_a^b (\alpha f_1 + \beta f_2)(t)\cos(x-t)dt \\ &= \int_a^b (\alpha f_1(t) + \beta f_2(t))\cos(x-t)dt \\ &= \int_a^b \alpha f_1(t)\cos(x-t)dt + \int_a^b \beta f_2(t)\cos(x-t)dt \\ &= \alpha \int_a^b f_1(t)\cos(x-t)dt + \beta \int_a^b f_2(t)\cos(x-t)dt \\ &= \alpha T(f_1)(x) + \beta T(f_2)(x); \end{aligned}$$

por tanto,

$$T(\alpha f_1 + \beta f_2) = \alpha T(f_1) + \beta T(f_2).$$

Lo cual prueba la linealidad de T.

(c) Sea $f \in C[a,b]$ y $g = T(f)$, entonces por (5.33)

$$g(x) = \alpha \cos x + \beta \operatorname{sen} x \quad \forall x \in [a,b]$$

donde $\alpha = \int_a^b f(t)\cos t\,dt$ y $\beta = \int_a^b f(t)\operatorname{sen} t\,dt$. Por tanto, $T(C[a,b]) = \operatorname{gn}(\cos x, \operatorname{sen} x)$; y ya que las funciones seno y coseno son L.I. en $C[a,b]$, se desprende que

$$\dim(T(C[a,b])) = 2.$$

(d) Puesto que $C[a,b]$ es un espacio vectorial de dimensión infinita y $\dim(T(C[a,b])) = 2$, del ejercicio 6 se concluye que la dimensión de $T(\mathbf{E})$ es infinita. ✔

8 Sean $\mathbf{E}$, $\mathbf{F}$ un par de espacios vectoriales. Probar que si $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$ y $V < \mathbf{F}$, entonces la imagen inversa de V bajo la transformación T, esto es

$$T^{-1}(V) = \{\vec{u} \in \mathbf{E} \,|\, T(\vec{u}) \in V\},$$

es un subespacio de $\mathbf{E}$.

DEMOSTRACIÓN ■ Dado que $V < \mathbf{F}$, $\vec{0}_{\mathbf{F}} \in V$ y puesto que $T(\vec{0}_{\mathbf{E}}) = \vec{0}_{\mathbf{F}}$, se tiene $\vec{0}_{\mathbf{E}} \in T^{-1}(V)$. Sean $\alpha, \beta \in \mathbb{R}$ y $\vec{u}_1, \vec{u}_2 \in T^{-1}(V)$, entonces, ya que $V < \mathbf{F}$, $T(\alpha\vec{u}_1 + \beta\vec{u}_2) = \alpha T(\vec{u}_1) + \beta T(\vec{u}_2) \in V$. Por tanto, $\alpha\vec{u}_1 + \beta\vec{u}_2 \in T^{-1}(V)$. Luego $T^{-1}(V) < \mathbf{E}$. ■

9 Sean $\mathbf{E}$ un espacio vectorial de dimensión finita n y $H < \mathbf{E}$. Se dice que H es un hiperplano de $\mathbf{E}$ si $\dim(H) = n - 1$. Probar que las siguientes condiciones son equivalentes a pares respecto a cualquier subespacio H de $\mathbf{E}$:

(a) H es un hiperplano de $\mathbf{E}$.

(b) Existe $f \in \mathscr{L}(\mathbf{E}, \mathbb{R})$, una forma lineal no nula, tal que $\operatorname{Ker}(f) = H$.

(c) Para toda base ordenada $(\vec{e}_1, \ldots, \vec{e}_n)$ de $\mathbf{E}$, existen escalares $\alpha_1, \ldots, \alpha_n$, no todos cero, tales que

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0$$

es una ecuación implícita de H; es decir, $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ es solución de esta ecuación si y sólo si $\vec{u} = \sum_{i=1}^{n} x_i \vec{e}_n \in H$.

DEMOSTRACIÓN ■ *(a)* $\Rightarrow$ *(b)* Sea $\{\vec{u}_1, \ldots, \vec{u}_{n-1}\}$ una base de H. Por el procedimiento dado en el apartado de la página 167 se puede completar ésta a una base $\{\vec{u}_1, \ldots, \vec{u}_{n-1}, \vec{u}_n\}$ del espacio $\mathbf{E}$. Por el ejercicio resuelto 1 existe una transformación lineal $f : \mathbf{E} \to \mathbb{R}$ tal que $f(\vec{u}_k) = 0$ para $k = 1, \ldots, n-1$ y $f(\vec{u}_n) = 1$. Entonces $f \neq \theta$, la transformación constante cero, y

$$\begin{aligned}
&\vec{u} = \sum_{i=1}^{n} a_i \vec{u}_i \in \operatorname{Ker}(f) \Leftrightarrow \\
&0 = f(\vec{u}) = \sum_{i=1}^{n-1} a_i f(\vec{u}_i) + a_n f(\vec{u}_n) \\
&\quad = a_n f(\vec{u}_n) \\
&\quad = a_n \Leftrightarrow \\
&\vec{u} \in H.
\end{aligned}$$

Luego $\operatorname{Ker}(f) = H$.

(b) $\Rightarrow$ *(c)* Sea $(\vec{e}_1, \ldots, \vec{e}_n)$ una base ordenada de $\mathbf{E}$. Sean $\alpha_i = f(\vec{e}_i)$, $i = 1, \ldots, n$; dado que f no es nula, alguno de los α_i debe ser distinto de cero. Si $\vec{u} = \sum_{i=1}^{n} x_i \vec{e}_i$, entonces

$$f(\vec{u}) = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n;$$

luego

$$\vec{u} \in \operatorname{Ker}(f) = H \Leftrightarrow \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0.$$

(c) $\Rightarrow$ *(a)* Sean $(\vec{e}_1, \ldots, \vec{e}_{n-1}, \vec{e}_n)$ una base ordenada de $\mathbf{E}$ y α_i, $i = 1, \ldots, n$, escalares no todos nulos tales que

$$\vec{u} = \sum_{i=1}^{n} x_i \vec{e}_i \in H \Leftrightarrow \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0.$$

Se puede suponer, sin perder generalidad, que $\alpha_n \neq 0$. Sea $\vec{u} = \sum_{i=1}^{n} x_i\vec{e}_i \in H$, entonces

$$x_n = -\frac{1}{\alpha_n}\sum_{i=1}^{n-1}\alpha_i x_i$$

y por tanto

$$\begin{aligned}\vec{u} &= x_1\vec{e}_1 + \cdots + x_{n-1}\vec{e}_{n-1} + x_n\vec{e}_n\\ &= x_1\vec{e}_1 + \cdots + x_{n-1}\vec{e}_{n-1} - \frac{1}{\alpha_n}\sum_{i=1}^{n-1}\alpha_i x_i\vec{e}_n\\ &= x_1\vec{e}_1 + \cdots + x_{n-1}\vec{e}_{n-1} - (\frac{\alpha_1}{\alpha_n}x_1 + \cdots + \frac{\alpha_{n-1}}{\alpha_n}x_{n-1})\vec{e}_n\\ &= x_1(\vec{e}_1 - \frac{\alpha_1}{\alpha_n}\vec{e}_n) + x_2(\vec{e}_2 - \frac{\alpha_2}{\alpha_n}\vec{e}_n) + \cdots + x_{n-1}(\vec{e}_{n-1} - \frac{\alpha_{n-1}}{\alpha_n}\vec{e}_n),\end{aligned}$$

luego

$$H = \operatorname{gn}(\vec{f}_1, \ldots, \vec{f}_{n-1})$$

donde

$$\vec{f}_k = \vec{e}_k - \frac{\alpha_k}{\alpha_n}\vec{e}_k$$

para $k = 1, \ldots, n-1$. Si $\beta_k \in \mathbb{R}$ son $n-1$ escalares tales que

$$\beta_1\vec{f}_1 + \cdots + \beta_{n-1}\vec{f}_{n-1} = \vec{0}_{\mathbf{E}},$$

entonces

$$\begin{aligned}\vec{0}_{\mathbf{E}} &= \beta_1\vec{e}_1 - \frac{\beta_1\alpha_1}{\alpha_n}\vec{e}_n + \beta_2\vec{e}_2 - \frac{\beta_2\alpha_2}{\alpha_n}\vec{e}_n + \cdots + \beta_{n-1}\vec{e}_{n-1} - \frac{\beta_{n-1}\alpha_{n-1}}{\alpha_n}\vec{e}_n\\ &= \beta_1\vec{e}_1 + \cdots + \beta_{n-1}\vec{e}_{n-1} + (-\frac{1}{\alpha_n}\sum_{i=1}^{n-1}\beta_i\alpha_i)\vec{e}_n\end{aligned}$$

y ya que los vectores $\vec{e}_i$, $i = 1, \ldots, n$, son L.I., se debe tener

$$\beta_1 = \cdots = \beta_{n-1} = 0$$

Por tanto los vectores $\vec{f}_1, \ldots, \vec{f}_{n-1}$ forman una base para H. Luego

$$\dim(H) = n-1$$

por ende H es un hiperplano de $\mathbf{E}$. ■

10 Demostrar que si $T : \mathbb{R}^n \to \mathbb{R}^n$ es un operador lineal invertible, $A \in \mathfrak{M}_n$ es su representación matricial relativa a la base canónica de $\mathbb{R}^n$ y T^{-1} es su operador inverso, entonces A es invertible y la representación matricial de T^{-1} relativa a la base canónica de $\mathbb{R}^n$ es A^{-1}.

DEMOSTRACIÓN ■ Por el teorema 5.9 A es invertible. Sea $\vec{v} \in \mathbb{R}^n$, puesto que T es suprayectivo (ya que es invertible), existe $\vec{u} \in \mathbb{R}^n$ tal que $T(\vec{u}) = \vec{v}$. Entonces

$$\begin{aligned} A^{-1}\vec{v} &= A^{-1}(A\vec{u}) \\ &= I_n\vec{u} \\ &= \vec{u} \\ &= T^{-1}(\vec{v}). \quad \blacksquare \end{aligned}$$

11 Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ la aplicación definida por

$$T(x,y) = (x-y, x+y)$$

(a) Probar que T es un operador lineal.

(b) Mostrar que T es invertible.

(c) Hallar T^{-1}.

DEMOSTRACIÓN ■ *(a)* $T(1,0) = (1,1)$, $T(0,1) = (-1,1)$. Si $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, entonces

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x-y \end{bmatrix} = T(x,y);$$

es decir, $T_A(x,y) = T(x,y)$ para todo $(x,y) \in \mathbb{R}^2$; por tanto, T es lineal y A es la representación matricial de T relativa a la base canónica de $\mathbb{R}^2$.

(b)

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix} \sim I_2$$

Por lo que A es invertible y por tanto T también (cfr. 5.9).

(c) Por el ejercicio anterior,

$$A^{-1} = -\frac{1}{2}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$$

es la representación matricial de T^{-1}; entonces,

$$\begin{aligned} T^{-1}(x,y) &= -\frac{1}{2}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2}x + \frac{1}{2}y \\ \frac{1}{2}x - \frac{1}{2}y \end{bmatrix} \end{aligned}$$

para todo $(x,y) \in \mathbb{R}^2$. ■

12 Sean T_1, T_2 un par de operadores lineales en un espacio vectorial $\mathbf{E}$. Se define la composición del operador T_2 con el operador T_1, $T_2 \circ T_1$, como la función $T_2 \circ T_1 : \mathbf{E} \to \mathbf{E}$ definida, para todo $\vec{u} \in \mathbf{E}$, por

$$(T_2 \circ T_1)(\vec{u}) = T_2(T_1(\vec{u})).$$

Probar que $T_2 \circ T_1$ es también un operador lineal. La composición $T_2 \circ T_1$ también se denota por T_2T_1; i.e., $T_2T_1 = T_2 \circ T_1$.

DEMOSTRACIÓN ■ Sean $\vec{u}, \vec{v} \in \mathbf{E}$ vectores arbitrarios y α, β cualquier par de escalares. Entonces

$$\begin{aligned}(T_2 \circ T_1)(\alpha\vec{u} + \beta\vec{v}) &= T_2(T_1(\alpha\vec{u} + \beta\vec{v}))\\ &= T_2(\alpha T_1(\vec{u}) + \beta T_1(\vec{v}))\\ &= \alpha T_2(T_1(\vec{u})) + \beta T_2(T_1(\vec{v}))\\ &= \alpha(T_2 \circ T_1)(\vec{u}) + \beta(T_2 \circ T_1)(\vec{v}). \quad ■\end{aligned}$$

13 Sean $T_1, T_2 : \mathbf{E} \to \mathbf{E}$ dos operadores lineales, demostrar lo siguiente.

(a) Si T_1 y T_2 son inyectivos, entonces T_2T_1 y T_1T_2 también son inyectivos.

(b) Si T_1T_2 no es inyectivo, entonces al menos uno de los operadores T_1, T_2 no es inyectivo.

(c) Si además T_1, T_2 son suprayectivos y uno de ellos (al menos) no es inyectivo, entonces T_1T_2 y T_2T_1 no son inyectivos.

DEMOSTRACIÓN ■ *(a)* Sea $\vec{u} \in \text{Ker}(T_2T_1)$, entonces

$$\begin{aligned}\vec{0}_{\mathbf{E}} &= (T_2T_1)(\vec{u})\\ &= T_2(T_1(\vec{u}))\end{aligned}$$

y puesto que T_2 es inyectivo, se debe tener $T_1(\vec{u}) = \vec{0}_{\mathbf{E}}$; y ya que este último operador también es inyectivo, se concluye $\vec{u} = \vec{0}_{\mathbf{E}}$ y, por tanto, $\text{Ker}(T_2T_1) = \{\vec{0}_{\mathbf{E}}\}$. Análogamente se demuestra la igualdad $\text{Ker}(T_1T_2) = \{\vec{0}_{\mathbf{E}}\}$, lo cual prueba que T_2T_1 y T_1T_2 son inyectivos.

(b) Si T_1 y T_2 son inyectivos, por el inciso anterior, T_1T_2 es inyectivo; lo cual es una contradicción. Por lo que uno de los dos operadores, por lo menos, no es inyectivo.

(c) Se prueba que T_1T_2 no es inyectivo, la demostración de que T_2T_1 no es inyectivo es entonces inmediata por simetría.

(i) Si T_2 no es inyectivo, existe $\vec{u} \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$ tal que $T_2(\vec{u}) = \vec{0}_{\mathbf{E}}$; entonces

$$\begin{aligned}(T_1T_2)(\vec{u}) &= T_1(T_2(\vec{u}))\\ &= T_1(\vec{0}_{\mathbf{E}})\\ &= \vec{0}_{\mathbf{E}}\end{aligned}$$

y por tanto T_1T_2 no es inyectivo.

(ii) Si T_1 no es inyectivo, existe $\vec{w} \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$ tal que $T_1(\vec{w}) = \vec{0}_{\mathbf{E}}$. Ya que T_2 es suprayectivo, existe $\vec{u} \in \mathbf{E}$ con $T_2(\vec{u}) = \vec{w}$; y como T_2 es lineal y $\vec{w} \neq \vec{0}_{\mathbf{E}}$, se desprende que $\vec{u} \neq \vec{0}_{\mathbf{E}}$. Entonces

$$\begin{aligned}(T_1T_2)(\vec{u}) &= T_1(T_2\vec{u})\\ &= T_1(\vec{w})\\ &= \vec{0}_{\mathbf{E}}.\end{aligned}$$

De donde se concluye que T_1T_2 no es inyectivo. ■

14 Sean $\mathbf{E}$ un espacio vectorial e $I : \mathbf{E} \to \mathbf{E}$ el operador lineal identidad; i.e., $I(\vec{u}) = \vec{u}$ para todo $\vec{u} \in \mathbf{E}$ (cfr. ejemplo 5.8). Probar que un operador lineal T en $\mathbf{E}$ es invertible si y sólo si existe un operador lineal T_1 en $\mathbf{E}$ tal que $T_1 \circ T = I = T \circ T_1$. En tal caso $T_1 = T^{-1}$.

DEMOSTRACIÓN ■ Si se supone que T es no singular (invertible), entonces claramente $T_1 = T^{-1}$ satisface la condición $T_1 \circ T = I = T \circ T_1$. Suponga que existe $T_1 \in \mathscr{L}(\mathbf{E},\mathbf{E})$ con $T_1 \circ T = I = T \circ T_1$. Si $\vec{u} \in \text{Ker}(T)$, entonces

$$\begin{aligned}\vec{u} &= I(\vec{u})\\ &= T_1(T(\vec{u}))\\ &= T_1(\vec{0}_{\mathbf{E}})\\ &= \vec{0}_{\mathbf{E}}\end{aligned}$$

y por tanto, $\text{Ker}(T) = \{\vec{0}_{\mathbf{E}}\}$; luego T es inyectiva. Si $\vec{v} \in \mathbf{E}$, sea $\vec{u} = T_1(\vec{v})$, entonces

$$\begin{aligned}T(\vec{u}) &= T(T_1(\vec{v}))\\ &= I(\vec{v}) = \vec{v}\end{aligned}$$

por ende T es suprayectiva. Así, T es biyectiva y por tanto invertible. Debido a la unicidad de T^{-1} se tiene $T^{-1} = T_1$. ■

15 Sean $\mathbf{E}$ un espacio vectorial y T un operador lineal en él. Se representan por T^2 el operador composición de T consigo mismo; i.e., $T^2 = T \circ T$ (cfr. el ejercicio resuelto 12 de esta sección) y por I el operador identidad, $I(\vec{u}) = \vec{u}$ para todo $\vec{u} \in \mathbf{E}$. Si $T^2 = I$ y $T \neq \pm I$ probar lo siguiente:

(a) Existe $\vec{v} \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$ tal que $T(\vec{v}) = \vec{v}$.

(b) Existe $\vec{w} \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$ tal que $T(\vec{w}) = -\vec{w}$.

DEMOSTRACIÓN ■ Dado que $T^2 = I$, se tiene $T(T(\vec{u})) = \vec{u}$ para todo $\vec{u} \in \mathbf{E}$.

(a) Puesto que $T \neq -I$, existe $\vec{u} \in \mathbf{E}$ tal que $T(\vec{u}) + \vec{u} \neq \vec{0}_{\mathbf{E}}$. Si $\vec{v} = T(\vec{u}) + \vec{u}$, entonces $\vec{v} \neq \vec{0}_{\mathbf{E}}$ y

$$\begin{aligned}T(\vec{v}) &= T(T(u) + \vec{u})\\ &= T(T(\vec{u})) + T(\vec{u})\\ &= \vec{u} + T(\vec{u})\\ &= \vec{v}.\end{aligned}$$

(b) Ya que $T \neq I$, existe $\vec{u} \in \mathbf{E}$ tal que $T(\vec{u}) - \vec{u} \neq \vec{0}_{\mathbf{E}}$. Sea $\vec{w} = T(\vec{u}) - \vec{u}$, entonces $\vec{w} \neq \vec{0}_{\mathbf{E}}$ y

$$\begin{aligned}T(\vec{w}) &= T(T(\vec{u}) - \vec{u})\\ &= T(T(\vec{u})) - T(\vec{u})\\ &= \vec{u} - T(\vec{u})\\ &= -\vec{w}.\end{aligned}$$ ■

16 Sean T el operador lineal del ejercicio anterior; esto es, $T^2 = I$ y $T \neq \pm I$ en el espacio $\mathbf{E}$;

$$\begin{aligned}S_1 &= \{\vec{v} \in \mathbf{E} \mid T(\vec{v}) = \vec{v}\};\\ S_2 &= \{\vec{v} \in \mathbf{E} \mid T(\vec{v}) = -\vec{v}\}.\end{aligned}$$

Probar que:

(a) S_1 y S_2 son subespacios de $\mathbf{E}$.

(b) $S_1 \neq \{\vec{0}_{\mathbf{E}}\} \neq S_2$.

(c) $\mathbf{E} = S_1 \oplus S_2$, la suma directa de S_1 con S_2 (cfr. ejercicio resuelto 27 del capítulo 3).

DEMOSTRACIÓN ■ *(a)* Ya que T es lineal $T(\vec{0}_{\mathbf{E}}) = \vec{0}_{\mathbf{E}}$, y por ello $\vec{0}_{\mathbf{E}} \in S_1 \cap S_2$. Sean $\alpha, \beta \in \mathbb{R}$ y $\vec{u}_1, \vec{v}_1 \in S_1$, $\vec{u}_2, \vec{v}_2 \in S_2$; entonces

$$T(\alpha\vec{u}_1 + \beta\vec{v}_1) = \alpha T(\vec{u}_1) + \beta T(\vec{v}_1) = \alpha\vec{u}_1 + \beta\vec{v}_1$$

y

$$T(\alpha\vec{u}_2 + \beta\vec{v}_2) = -\alpha T(\vec{u}_2) - \beta T(\vec{v}_2) = -(\alpha\vec{u}_2 + \beta\vec{v}_2).$$

Lo cual prueba $S_1 < \mathbf{E}$ y $S_2 < \mathbf{E}$.

(b) Es consecuencia inmediata del ejercicio precedente.

(c) Si $\vec{u} \in S_1 \cap S_2$, entonces

$$\vec{u} = T(\vec{u}) = -\vec{u};$$

de donde $\vec{u} = \vec{0}_{\mathbf{E}}$. Sea $\vec{v} \in \mathbf{E}$ y $\vec{u}_1 = \frac{1}{2}(\vec{v} + T(\vec{v}))$, $\vec{u}_2 = \frac{1}{2}(\vec{v} - T(\vec{v}))$. Entonces

$$T(\vec{u}_1) = \frac{1}{2}(T(\vec{v}) + \vec{v}) = \vec{u}_1,$$
$$T(\vec{u}_2) = \frac{1}{2}(T(\vec{v}) - \vec{v}) = -\vec{u}_2;$$

por tanto $\vec{u}_1 \in S_1$ y $\vec{u}_2 \in S_2$. Y ya que

$$\vec{v} = \vec{u}_1 + \vec{u}_2$$

se desprende que $\mathbf{E} = S_1 \oplus S_2$. ■

17 Sean $\mathbf{E}$ un espacio vectorial y $S_1 < \mathbf{E}$. Una *proyección* sobre S_1 es una función $T : \mathbf{E} \to \mathbf{E}$ tal que:

(i) Existe $S_2 < \mathbf{E}$ que satisface $\mathbf{E} = S_1 \oplus S_2$.

(ii) Para todo $\vec{u} = \vec{x}_1 + \vec{x}_2$, con $\vec{x}_i \in S_i$, $i = 1, 2$, se tiene $T(\vec{u}) = \vec{x}_1$.

Mostrar que si T es una proyección sobre S_1, entonces:

(a) T es lineal.

(b) $S_1 = \{\vec{u} \in \mathbf{E} \,|\, T(\vec{u}) = \vec{u}\}$.

(c) $S_1 = T(\mathbf{E})$.

(d) $S_2 = \mathrm{Ker}(T)$.

DEMOSTRACIÓN ■ *(a)* Si $\alpha, \beta \in \mathbb{R}$ y $\vec{u} = \vec{x}_1 + \vec{x}_2 \in \mathbf{E}$, $\vec{v} = \vec{y}_1 + \vec{y}_2 \in \mathbf{E}$, con $\vec{x}_i \in S_1$ y $\vec{y}_i \in S_2$, $i = 1, 2$; entonces

$$\begin{aligned} T(\alpha\vec{u} + \beta\vec{v}) &= T((\alpha\vec{x}_1 + \beta\vec{y}_1) + (\alpha\vec{x}_2 + \beta\vec{y}_2)) \\ &= \alpha\vec{x}_1 + \beta\vec{y}_1 \\ &= \alpha T(\vec{u}) + \beta T(\vec{v}). \end{aligned}$$

(b) $\vec{x}_1 \in S_1 \Rightarrow \vec{x}_1 = \vec{x}_1 + \vec{0}_{\mathbf{E}} \Rightarrow T(\vec{x}_1) = T(\vec{x}_1 + \vec{0}_{\mathbf{E}}) = \vec{x}_1$. $\vec{u} = \vec{x}_1 + \vec{x}_2, \vec{u} = T(\vec{u}) \Rightarrow \vec{x}_1 + \vec{x}_2 = \vec{x}_1 \Rightarrow \vec{u} \in S_1$.

(c) Si $\vec{x}_1 \in S_1$, entonces $\vec{x}_1 = T(\vec{x}_1) \in T(\mathbf{E})$ y por tanto $S_1 \subset T(S_1)$. Si $T(\vec{x}_1 + \vec{x}_2) \in T(\mathbf{E})$, entonces $T(\vec{x}_1 + \vec{x}_2) = \vec{x}_1 \in S_1$ y por ende $T(\mathbf{E}) \subset S_1$. Luego $T(\mathbf{E}) = S_1$.

(d) $\vec{u} = \vec{x}_1 + \vec{x}_2 \in \operatorname{Ker}(T) \Rightarrow \vec{x}_1 = T(\vec{x}_1 + \vec{x}_2) = \vec{0}_{\mathbf{E}} \Rightarrow \vec{u} = \vec{x}_2 \in S_2$; $\therefore \operatorname{Ker}(T) \subset S_2$. $\vec{x}_2 \in S_2 \Rightarrow \vec{x}_2 = \vec{0}_{\mathbf{E}} + \vec{x}_2 \Rightarrow T(\vec{x}_2) = \vec{0}_{\mathbf{E}} \Rightarrow S_2 \subset \operatorname{Ker}(T)$. Luego $\operatorname{Ker}(T) = S_2$. ■

18 Sean $\mathbf{E}$ un espacio vectorial, T un operador lineal en él y $S = \{\vec{u} \in \mathbf{E} \,|\, T(\vec{u}) = \vec{u}\}$; por el ejercicio anterior S es un subespacio de $\mathbf{E}$.

(a) Si $T^2 = T$, demostrar que $\mathbf{E} = S \oplus \operatorname{Ker}(T)$.

(b) Inversamente, si $\mathbf{E} = S \oplus \operatorname{Ker}(T)$, probar que $T^2 = T$.

DEMOSTRACIÓN ■ *(a)* Si $\vec{u} \in \mathbf{E}$, entonces

$$\vec{u} = T(\vec{u}) + (\vec{u} - T(\vec{u}))$$

y

$$\begin{aligned} T(T(\vec{u})) &= T^2(\vec{u}) = T(\vec{u}), \\ T(\vec{u} - T(\vec{u})) &= T(\vec{u}) - T^2(\vec{u}) = \vec{0}_{\mathbf{E}}. \end{aligned}$$

Por tanto, $T(\vec{u}) \in S$, $\vec{u} - T(\vec{u}) \in \operatorname{Ker}(T)$ y $\mathbf{E} = S + \operatorname{Ker}(T)$. Si $\vec{u} \in S \cap \operatorname{Ker}(T)$, $\vec{u} = T(\vec{u}) = \vec{0}_{\mathbf{E}}$; luego $S \cap \operatorname{Ker}(T) = \{\vec{0}_{\mathbf{E}}\}$.

(b) Sea $\vec{u} \in \mathbf{E}$, entonces existen $\vec{u}_1 \in S$ y $\vec{u}_2 \in \operatorname{Ker}(T)$ tales que $\vec{u} = \vec{u}_1 + \vec{u}_2$; por lo que $T(\vec{u}) = T(\vec{u}_1) + T(\vec{u}_2) = \vec{u}_1$ y entonces $T^2(\vec{u}) = T(T(\vec{u})) = T(\vec{u}_1) = T(\vec{u})$. ■

Representaciones matriciales de transformaciones lineales

19 Sean las bases $\mathscr{B}_1 = \{(1, -1), (2, -1)\}$ y $\mathscr{B}_2 = \{(-1, 1), (0, 1)\}$ del espacio $\mathbb{R}^2$.

(a) Encontrar la matriz cambio de base de la base $\mathscr{B}_2$ a la base $\mathscr{B}_1$.

(b) Encontrar la matriz cambio de base de la base $\mathscr{B}_1$ a la base $\mathscr{B}_2$.

Solución *(a)* Se tienen que resolver los sistemas con la misma matriz de coeficientes

$$\begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

y

$$\begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Dado que

$$\left[\begin{array}{cc|cc} 1 & 2 & -1 & 0 \\ -1 & -1 & 1 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 2 & -1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & -1 & -2 \\ 0 & 1 & 0 & 1 \end{array}\right]$$

se tiene

$$\begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix};$$

esto es,

$$P = \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix}$$

es la matriz cambio de base de la base $\mathscr{B}_2$ a la base $\mathscr{B}_1$.

(b) La matriz cambio de base de $\mathscr{B}_1$ a $\mathscr{B}_2$ está dada por

$$P^{-1} = -1\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix}. \quad ✔$$

20 Sean T_1, T_2 operadores lineales en un espacio vectorial **E** que tiene dimensión finita y $\mathscr{B}$ una base de **E**. Si $T = T_2 \circ T_1$ es el operador composición definido en el ejercicio resuelto 12 de este apartado y $[T_1]_{\mathscr{B}}$, $[T_2]_{\mathscr{B}}$ son las representaciones matriciales de estos operadores relativas a la base $\mathscr{B}$, mostrar que

$$[T_2 \circ T_1]_{\mathscr{B}} = [T_2]_{\mathscr{B}}[T_1]_{\mathscr{B}}$$

o, con la notación $T_2T_1 = T_2 \circ T_1$,

$$[T_2T_1]_{\mathscr{B}} = [T_2][T_1]_{\mathscr{B}}.$$

DEMOSTRACIÓN ■ Si $\vec{u} \in \mathbf{E}$ cualquier vector, entonces

$$\begin{aligned} [(T_2 \circ T_1)\vec{u}]_{\mathscr{B}} &= [T_2(T_1(\vec{u}))]_{\mathscr{B}} \\ &= [T_2]_{\mathscr{B}}[T_1(\vec{u})]_{\mathscr{B}} \\ &= [T_2]_{\mathscr{B}}([T_1]_{\mathscr{B}}[\vec{u}]_{\mathscr{B}}) \\ &= ([T_2]_{\mathscr{B}}[T_1]_{\mathscr{B}})[\vec{u}]_{\mathscr{B}}. \end{aligned}$$

De donde se desprende, debido a la unicidad de la representación matricial de un operador, que

$$[T_2 \circ T_1]_{\mathscr{B}} = [T_2]_{\mathscr{B}}[T_1]_{\mathscr{B}}. \quad ■$$

21 Sean **E** un espacio de dimensión finita dotado de un producto interior $\langle \cdot, \cdot \rangle$, $\mathscr{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ una base ortonormal de este espacio y $T : \mathbf{E} \to \mathbf{E}$ un operador lineal. Encontrar la representación matricial de T relativa a la base $\mathscr{B}$.

DEMOSTRACIÓN ■ Si $a_{ij} \in \mathbb{R}$ son tales que

$$T(\vec{e}_i) = a_{i1}\vec{e}_1 + \cdots + a_{ij}\vec{e}_i + \cdots + a_{in}\vec{e}_n$$

entonces, para cada $j = 1, \ldots, n$,

$$\begin{aligned} \langle T(\vec{e}_i), \vec{e}_j \rangle &= \langle a_{i1}\vec{e}_1 + \cdots + a_{ij}\vec{e}_i + \cdots + a_{in}\vec{e}_n, \vec{e}_j \rangle \\ &= a_{ij} \end{aligned}$$

y por tanto

$$[T]_{\mathscr{B}} = \begin{bmatrix} \langle T(\vec{e}_1), \vec{e}_1 \rangle & \langle T(\vec{e}_2), \vec{e}_1 \rangle & \cdots & \langle T(\vec{e}_n), \vec{e}_1 \rangle \\ \langle T(\vec{e}_1), \vec{e}_2 \rangle & \langle T(\vec{e}_2), \vec{e}_2 \rangle & \cdots & \langle T(\vec{e}_n), \vec{e}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle T(\vec{e}_1), \vec{e}_n \rangle & \langle T(\vec{e}_2), \vec{e}_n \rangle & \cdots & \langle T(\vec{e}_n), \vec{e}_n \rangle \end{bmatrix}. \quad ■$$

22 Sean **E** un espacio vectorial y H_n una sucesión de subespacios tales que $H_n \subsetneq H_{n+1}$ para todo n. Probar que existe una sucesión infinita $(\vec{u}_n)$ de vectores L.I. en **E** y que por tanto **E** tiene dimensión infinita.

DEMOSTRACIÓN ■ Se puede suponer, sin perder generalidad, que $H_1 \neq \{\vec{0}_\mathbf{E}\}$. Entonces existe $\vec{u} \in H_1 - \{\vec{0}_\mathbf{E}\}$; sea $\vec{u}_1 = \vec{u}$. Puesto que $H_{n-1} \neq H_n$, existe, para cada $n = 2, 3, \ldots$, $\vec{u}_n \in H_n - H_{n-1}$. La sucesión de vectores $(\vec{u}_n)$ así formada es L.I. Para probar esto basta mostrar que para todo n el conjunto finito $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ es L.I. Se procede por inducción sobre n: si $n = 1$, $\{\vec{u}_1\}$ es L.I. porque $\vec{u}_1 \neq \vec{0}_\mathbf{E}$. Sea $k > 1$ y suponga que la afirmación es cierta para $n = k - 1$. Sean α_i, $i = 1, \ldots, k$, escalares tales que

$$\alpha_1 u_1 + \cdots + \alpha_{k-1}\vec{u}_{k-1} + \alpha_k \vec{u}_k = \vec{0}_\mathbf{E}.$$

Si $\alpha_k \neq 0$, entonces

$$\vec{u}_k = -\frac{\alpha_1}{\alpha_k}\vec{u}_1 - \cdots - \frac{\alpha_{k-1}}{\alpha_k}\vec{u}_{k-1} \in H_{k-1}$$

pues $H_1 \subset H_2 \subset \cdots \subset H_{k-1}$; por tanto, α_k debe ser cero. Luego

$$\alpha_1 u_1 + \cdots + \alpha_{k-1}\vec{u}_{k-1} = \vec{0}_\mathbf{E}$$

y por la hipótesis de inducción se concluye que

$$\alpha_1 = \cdots = \alpha_{k-1} = \alpha_k = 0;$$

por ende, $\{\vec{u}_1, \ldots, \vec{u}_{k-1}, \vec{u}_k\}$ es L.I. ■

En los ejercicios 23 a 27 **E** es un espacio vectorial y $T : \mathbf{E} \to \mathbf{E}$ un operador lineal en el espacio **E**. Se define $T^2 = T \circ T$ y, por recurrencia, $T^n = T \circ T^{n-1}$, $n = 2, 3, \ldots$, esto es (cfr. ejercicio resuelto 12 de esta sección),

$$T^n = \underbrace{T \circ T \circ \cdots \circ T}_{n}$$

y si I es el operador lineal identidad, $I(\vec{u}) = \vec{u}$ para todo $\vec{u} \in \mathbf{E}$, se define $T^0 = I$. Suponga que $\mathbf{E}$ tiene dimensión finita.

23 Probar que $\operatorname{Ker}(T^n) \subset \operatorname{Ker}(T^{n+1})$ para todo $n = 0, 1, 2, \ldots$

DEMOSTRACIÓN ■

$$\begin{aligned}\vec{u} \in \operatorname{Ker}(T^n) &\Rightarrow T^n(\vec{u}) = \vec{0}_{\mathbf{E}}\\ &\Rightarrow T(T^n(\vec{u})) = \vec{0}_{\mathbf{E}}\\ &\Rightarrow T^{n+1}(\vec{u}) = \vec{0}_{\mathbf{E}}\\ &\Rightarrow \vec{u} \in \operatorname{Ker}(T^{n+1})\\ &\Rightarrow \operatorname{Ker}(T^n) \subset \operatorname{Ker}(T^{n+1}). \quad ■\end{aligned}$$

24 Mostrar que $T^{n+1}(\mathbf{E}) \subset T^n(\mathbf{E})$ para todo $n = 0, 1, 2, \ldots$

DEMOSTRACIÓN ■ Si $\vec{w} \in T^{n+1}(\mathbf{E})$, existe $\vec{u} \in \mathbf{E}$ tal que $T^{n+1}(\vec{u}) = \vec{w}$ y por tanto $T^n(T(\vec{u})) = \vec{w}$; luego $\vec{w} \in T^n(\mathbf{E})$. Lo cual prueba $T^{n+1}(\mathbf{E}) \subset T^n(\mathbf{E})$. ■

25 Mostrar que existe n_0 tal que:

(a) Si $n \geq n_0$:
- *(i)* $\operatorname{Ker}(T^n) = \operatorname{Ker}(T^{n+1})$.
- *(ii)* $T^{n+1}(\mathbf{E}) = T^n(\mathbf{E})$.

(b) Si $n < n_0$:
- *(i)* $\operatorname{Ker}(T^n) \neq \operatorname{Ker}(T^{n+1})$.
- *(ii)* $T^{n+1}(\mathbf{E}) \neq T^n(\mathbf{E})$.

DEMOSTRACIÓN ■ Si $\operatorname{Ker}(T^n) \subsetneq \operatorname{Ker}(T^{n+1})$ para todo n, entonces, por ejercicio resuelto 22, $\mathbf{E}$ tendría dimensión infinita; por tanto, deben existir enteros no negativos n tales que $\operatorname{Ker}(T^n) = \operatorname{Ker}(T^{n+1})$; sea n_0 el menor de éstos.

(a) Sea n un entero con $n > n_0$.
- *(i)* Si $\vec{u} \in \operatorname{Ker}(T^{n+1})$, entonces

$$\begin{aligned}\vec{0}_{\mathbf{E}} &= T^{n+1}(\vec{u})\\ &= T^{n-n_0+n_0+1}(\vec{u})\\ &= T^{n_0+1}(T^{n-n_0}(\vec{u}))\end{aligned}$$

y por tanto

$$T^{n-n_0}(\vec{u}) \in \operatorname{Ker}(T^{n_0+1}) = \operatorname{Ker}(T^{n_0})$$

luego

$$T^n(\vec{u}) = T^{n_0}(T^{n-n_0}(\vec{u})) = \vec{0}_{\mathbf{E}},$$

por lo que

$$\vec{u} \in \operatorname{Ker}(T^n).$$

Así que

$$\operatorname{Ker}(T^n) = \operatorname{Ker}(T^{n+1}).$$

(ii) Puesto que $\operatorname{Ker}(T^{n+1}) = \operatorname{Ker}(T^n)$,

$$\begin{aligned}\dim(\operatorname{Ker}(T^{n+1})) + \dim(T^{n+1}(\mathbf{E})) &= \dim(\mathbf{E})\\ &= \dim(\operatorname{Ker}(T^n)) + \dim(T^n(\mathbf{E}));\end{aligned}$$

implica

$$\dim(T^n(\mathbf{E})) = \dim(T^{n+1}(\mathbf{E}))$$

y ya que $T^{n+1}(\mathbf{E}) < T^n(\mathbf{E})$, se concluye que

$$T^{n+1}(\mathbf{E}) = T^n(\mathbf{E}).$$

(b) Sea n_0 un entero no negativo con $n < n_0$.

(i) Por definición de n_0

$$\operatorname{Ker}(T^n) \neq \operatorname{Ker}(T^{n+1}).$$

(ii) Por el inciso precedente, se concluye que $\dim(T^{n+1}(\mathbf{E})) \neq \dim(T^n(\mathbf{E}))$, y por ende,

$$T^{n+1}(\mathbf{E}) \neq T^n(\mathbf{E}). \quad ■$$

26 Probar que si $n \geq n_0$, entonces (cfr. el ejercicio resuelto 27 del capítulo 3, pág. 187).

$$\mathbf{E} = \operatorname{Ker}(T^n) \oplus T^n(\mathbf{E}).$$

DEMOSTRACIÓN ■ Si $\vec{u} \in \operatorname{Ker}(T^n) \cap T^n(\mathbf{E})$, existe $\vec{w} \in \mathbf{E}$ tal que $T^n(\vec{w}) = \vec{u}$ y por ello

$$\begin{aligned}\vec{0}_{\mathbf{E}} &= T^n(\vec{u})\\ &= T^{2n}(\vec{w});\end{aligned}$$

por tanto,

$$\vec{w} \in \operatorname{Ker}(T^{2n}) = \operatorname{Ker}(T^n),$$

entonces

$$\vec{u} = T^n(\vec{w}) = \vec{0}_{\mathbf{E}}.$$

Puesto que

$$\begin{aligned}\dim(\operatorname{Ker}(T^n)+T^n(\mathbf{E})) &= \dim(\operatorname{Ker}(T^n))+\dim(T^n(\mathbf{E}))-\dim(\operatorname{Ker}(T^n)\cap T^n(\mathbf{E}))\\ &= \dim(\operatorname{Ker}(T^n))+\dim(T^n(\mathbf{E})),\end{aligned}$$

se tiene

$$\begin{aligned}\dim(\mathbf{E}) &= \dim(\operatorname{Ker}(T^n))+\dim(T^n(\mathbf{E}))\\ &= \dim(\operatorname{Ker}(T^n)+\dim(T^n(\mathbf{E}))\end{aligned}$$

y por tanto,

$$\mathbf{E}=\operatorname{Ker}(T^n)+T^n(\mathbf{E}).$$

Lo cual prueba que

$$\mathbf{E}=\operatorname{Ker}(T^n)\oplus T^n(\mathbf{E}). \quad ■$$

27 Encontrar n_0 del ejercicio precedente para el operador lineal $T:\mathbb{R}^4\to\mathbb{R}^4$ definido por

$$T(x_1,x_2,x_3,x_4)=(-x_1-x_2,x_1+x_2,-x_1-2x_2-x_3+x_4,-x_1-2x_2-3x_3+3x_4).$$

Solución Puesto que

$$T(1,0,0,0)=(-1,1,-1,-1),$$

$$T(0,1,0,0)=(-1,1,-2,-2),$$

$$T(0,0,1,0)=(0,0,-1,-3) \text{ y}$$

$$T(0,0,0,1)=(0,0,1,3)$$

la representación matricial relativa a la base canónica de $\mathbb{R}^4$ para T es

$$A=\begin{bmatrix}-1 & -1 & 0 & 0\\ 1 & 1 & 0 & 0\\ -1 & -2 & -1 & 1\\ -1 & -2 & -3 & 3\end{bmatrix}.$$

Ya que, al hacer operaciones entre columnas,

$$A\sim\begin{bmatrix}-1 & -1 & 0 & 0\\ 1 & 1 & 0 & 0\\ -1 & -2 & -1 & 0\\ -1 & -2 & -3 & 0\end{bmatrix}$$

se deduce que $\operatorname{Rang}(A)=3$; por tanto, $\dim(\operatorname{Ker}(T))=4-3=1$. En el ejercicio resuelto 20 de este apartado se probó que la representación matricial de T^2es A^2; y por ello la representación matricial de T^n es A^n. Entonces, ya que

$$A^2 = \begin{bmatrix} -1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & -2 & -1 & 1 \\ -1 & -2 & -3 & 3 \end{bmatrix}^2$$
$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & -2 & 2 \\ -1 & -1 & -6 & 6 \end{bmatrix}$$
$$\sim \begin{bmatrix} -1 & -1 & -2 & 2 \\ 0 & 0 & -4 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 4 \\ 0 & 0 & -12 & 12 \end{bmatrix}$$
$$\sim \begin{bmatrix} 0 & 0 & -4 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

y

$$A^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -8 & 8 \\ 0 & 0 & -24 & 24 \end{bmatrix}$$
$$\sim \begin{bmatrix} 0 & 0 & -8 & 8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

se tiene

$$\dim(\text{Ker}(T^2)) = 2,$$
$$\dim(\text{Ker}(T^3)) = 3,$$
$$\dim(\text{Ker}(T^4)) = 3.$$

Por lo que $n_0 = 3$. ✔

28 Sean $\mathbf{E}$ un espacio vectorial, $\mathscr{B} = (\vec{e}_1, \ldots, \vec{e}_n)$ una base de él y T un operador lineal en $\mathbf{E}$. Para cada $k = 1, \ldots, n$ sea $S_k = \text{gn}(\vec{e}_1, \ldots, \vec{e}_k)$. Mostrar que las siguientes condiciones son equivalentes.

(a) $[T]_{\mathscr{B}}$ es triangular superior.

(b) Para cada $k = 1, \ldots, n$, $T(\vec{e}_i) \in S_k$ para todo $i = 1, \ldots, k$.

(c) $T(S_k) \subset S_k$ para todo $k = 1, \ldots, n$.

DEMOSTRACIÓN ■ *(a)* $\Rightarrow$ *(b)*: Si $[T]_{\mathscr{B}} = [a_{ij}]$ es triangular superior, entonces $T(\vec{e}_k) = \sum_{i=1}^{k} a_{ik}\vec{e}_i \in S_k$ para todo $k = 1,\ldots,n$.

(b) $\Rightarrow$ *(c)*: Si $\vec{u} \in S_k$, existen $\alpha_i \in \mathbb{R}$ tales que $\vec{u} = \sum_{i=1}^{k} \alpha_i\vec{e}_i$ y por tanto $T(\vec{u}) = \sum_{i=1}^{k} \alpha_i T(\vec{e}_i) \in S_k$.

(c) $\Rightarrow$ *(a)*: Puesto que $T(\vec{e}_k) \in S_k$, existen $a_{1k},\ldots,a_{kk} \in \mathbb{R}$ tales que $T(\vec{e}_k) = \sum_{i=1}^{k} a_{ki}\vec{e}_i$; de donde $[T]_{\mathscr{B}} = [a_{ij}]$ es triangular superior. ■

29 Sean $\mathbf{E}$ un espacio vectorial y T un operador lineal en él tal que $T^2 = -I$.

(a) Demostrar que T es invertible.

(b) Sean $\vec{u}_1,\ldots,\vec{u}_m \in \mathbf{E}$ tales que $\vec{u}_1,\ldots,\vec{u}_m$, $T(\vec{u}_1),\ldots,T(\vec{u}_{m-1})$ son L.I., probar que $\vec{u}_1,\ldots,\vec{u}_m$, $T(\vec{u}_1),\ldots,T(\vec{u}_{m-1}),T(\vec{u}_m)$ son L.I.

(c) Si $\mathbf{E}$ tiene dimensión finita distinta de cero, mostrar que $\mathbf{E}$ tiene una base de la forma $\mathscr{B} = \{\vec{u}_1,\ldots,\vec{u}_m,T(\vec{u}_1),\ldots,T(\vec{u}_{m-1}),T(\vec{u}_m)\}$ y por tanto su dimensión par.

(d) Encontrar la representación matricial de T relativa a la base $\mathscr{B}$ del inciso anterior.

DEMOSTRACIÓN ■ *(a)* $T(-T) = -T^2 = -(-I) = I$; $(-T)T = -T^2 = -(-I) = I$. Así T es invertible y $T^{-1} = -T$.

(b) Si $\vec{u}_1,\ldots,\vec{u}_m,T(\vec{u}_1),\ldots,T(\vec{u}_m)$ son L.D., entonces existen escalares α_i,β_i, $i = 1,\ldots,m$, con alguno de ellos distinto de cero, tales que

$$\vec{0}_{\mathbf{E}} = \sum_{i=1}^{m} \alpha_i\vec{u}_i + \sum_{i=1}^{m} \beta_i T(\vec{u}_i).$$

β_m debe ser distinto de cero; pues en caso contrario los vectores $\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_m,T(\vec{u}_1),\ldots,T(\vec{u}_{m-1})$ serían L.D. Entonces existen escalares γ_i y δ_j, $i = 1,\ldots,m$, $j = 1,\ldots,m-1$, tales que

$$T(\vec{u}_m) = \sum_{i=1}^{m} \gamma_i\vec{u}_i + \sum_{i=1}^{m-1} \delta_i T(\vec{u}_i)$$

por lo que

$$\begin{aligned} -\vec{u}_m &= T^2(\vec{u}_m) \\ &= TT(\vec{u}_m) \\ &= \sum_{i=1}^{m} \gamma_i T(\vec{u}_i) - \sum_{i=1}^{m-1} \delta_i\vec{u}_i, \end{aligned}$$

de donde

$$\begin{aligned} \vec{0}_{\mathbf{E}} &= \sum_{i=1}^{m} \gamma_i T(\vec{u}_i) - \sum_{i=1}^{m-1} \delta_i\vec{u}_i + \vec{u}_m \\ &= \sum_{i=1}^{m-1} \gamma_i T(\vec{u}_i) + \gamma_m T(\vec{u}_m) - \sum_{i=1}^{m-1} \delta_i\vec{u}_i + \vec{u}_m \\ &= \sum_{i=1}^{m-1} \gamma_i T(\vec{u}_i) + \gamma_m \left(\sum_{i=1}^{m} \gamma_i\vec{u}_i + \sum_{i=1}^{m-1} \delta_i T(\vec{u}_i) \right) - \sum_{i=1}^{m-1} \delta_i\vec{u}_i + \vec{u}_m \\ &= \sum_{i=1}^{m-1} (\gamma_i + \gamma_m\delta_i) T(\vec{u}_i) + \sum_{i=1}^{m-1} (\gamma_m\gamma_i - \delta_i)\vec{u}_i + (\gamma_m^2 + 1)\vec{u}_m. \end{aligned}$$

Lo cual implica, puesto que $\gamma_m^2+1\neq 0$, que los vectores $\vec{u}_1,\ldots,\vec{u}_m$, $T(\vec{u}_1),\ldots,T(\vec{u}_{m-1})$ son L.D., que es una contradicción de la hipótesis inicial. Luego los vectores $\vec{u}_1,\ldots,\vec{u}_m$, $T(\vec{u}_1),\ldots,T(\vec{u}_m)$ deben ser L.I.

(c) Sea $\vec{u}_1\in\mathbf{E}-\{\vec{0}_\mathbf{E}\}$. Si existe $k\in\mathbb{R}$ tal que $T(\vec{u}_1)=k\vec{u}_1$, entonces

$$\begin{aligned}-u_1 &= T^2(\vec{u}_1)\\ &= T(T(\vec{u}_1))\\ &= T(k\vec{u}_1)\\ &= kT(\vec{u}_1)\\ &= k^2\vec{u}_1\end{aligned}$$

luego

$$(k^2+1)\vec{u}_1=\vec{0}_\mathbf{E};$$

lo cual es imposible pues $\vec{u}_1\neq\vec{0}_\mathbf{E}$. Por tanto, $\vec{u}_1$ y $T(\vec{u}_1)$ son L.I. Si no existe un vector $\vec{u}_2\in\mathbf{E}$ tal que $\{\vec{u}_1,T(\vec{u}_1),\vec{u}_2\}$ es L.I., entonces $\{\vec{u}_1,T(\vec{u}_1)\}$ es una base para $\mathbf{E}$ y $\dim(\mathbf{E})=2$. Si existe $\vec{u}_2\in\mathbf{E}$ tal que $\vec{u}_1,T(\vec{u}_1),\vec{u}_2$ son L.I., entonces, por el inciso anterior, $\vec{u}_1,\vec{u}_2,T(\vec{u}_1),T(\vec{u}_2)$ son L.I. Si no existe $\vec{u}_3\in\mathbf{E}$ tal que $\vec{u}_1,\vec{u}_2,T(\vec{u}_1),T(\vec{u}_2),\vec{u}_3$ son L.I., entonces $\{\vec{u}_1,\vec{u}_2,T(\vec{u}_1),T(\vec{u}_2)\}$ es una base de $\mathbf{E}$ y $\dim(\mathbf{E})=4$. Continuando este proceso se debe llegar a un primer número entero m tal que

$$\vec{u}_1,\ldots,\vec{u}_m,T(\vec{u}_1),\ldots,T(\vec{u}_m)$$

son L.I. y para cualquier vector $\vec{u}\in\mathbf{E}$ el conjunto

$$\{\vec{u}_1,\ldots,\vec{u}_m,T(\vec{u}_1),\ldots,T(\vec{u}_m),\vec{u}\}$$

es L.D.; pues $\mathbf{E}$ tiene dimensión finita. Luego

$$\{\vec{u}_1,\ldots,\vec{u}_m,T(\vec{u}_1),\ldots,T(\vec{u}_m)\}$$

es una base de $\mathbf{E}$ y $\dim(\mathbf{E})=2m$.

(d) Si $\mathscr{B}=\{\vec{u}_1,\ldots,\vec{u}_m,T(\vec{u}_1),\ldots,T(\vec{u}_m)\}$, entonces

$$T(T(\vec{u}_i))=-u_i \text{ para } i=1,\ldots,m$$

luego

$$[T]_\mathscr{B}=\begin{bmatrix}\mathscr{O} & -I_m\\ I_m & \mathscr{O}\end{bmatrix}$$

donde la submatriz I_m es la identidad de orden m y la submatriz $\mathscr{O}$ es la matriz cero de orden $m\times m$. ■

30 (**Espacio dual y base dual**). Sea $\mathbf{E}$ un espacio vectorial; se denota por $\mathbf{E}^*$ el espacio vectorial $\mathscr{L}(\mathbf{E},\mathbb{R})$ y a toda transformación lineal $f\in\mathbf{E}^*$ se le dice *funcional lineal* (o *forma lineal*) en $\mathbf{E}$ y a $\mathbf{E}^*$ se le llama el espacio *dual* de $\mathbf{E}$. Si $\mathbf{E}$ tiene dimensión finita y $\mathscr{B}=\{\vec{e}_1,\ldots,\vec{e}_n\}$ es una base de $\mathbf{E}$, encontrar una base $\mathscr{B}^*=\{\phi_i\}$ del dual del espacio $\mathbf{E}$; esto es, una base de $\mathbf{E}^*$ a partir de esta base y probar que $\dim(\mathbf{E}^*)=n$ y por tanto $\mathbf{E}\cong\mathbf{E}^*$. La base $\mathscr{B}^*=\{\phi_i\}$ se llama base dual de la base $\{\vec{e}_i\}$.

Solución Sea $i \in \{1, \dots, n\}$, por el ejercicio resuelto 1 existe un único funcional lineal $\phi_i : \mathbf{E} \to \mathbb{R}$ tal que

$$\phi_i(\vec{e}_j) = \delta_{ij} = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

(al símbolo δ_{ij} se le dice *delta de Kronecker*). Entonces $\{\phi_1, \phi_2, \dots, \phi_n\}$ es una base de $\mathbf{E}^*$:

(a) Sean α_i n-escalares tales que

$$\alpha_1\phi_1 + \cdots + \alpha_i\phi_i + \cdots + \alpha_n\phi_n = \theta,$$

la forma lineal constante cero en $\mathbf{E}^*$; entonces, para cada $i = 1, \dots, n$,

$$\begin{aligned} \alpha_i &= \alpha_1\phi_1(\vec{e}_i) + \cdots + \alpha_i\phi_i(\vec{e}_i) + \cdots + \alpha_n\phi_n(\vec{e}_i) \\ &= \theta(\vec{e}_i) \\ &= 0 \end{aligned}$$

así que $\{\phi_1, \phi_2, \dots, \phi_n\}$ es L.I.

(b) Sea $f \in \mathbf{E}^*$ y $\vec{u} = a_1\vec{e}_1 + \cdots + a_i\vec{e}_i + \cdots + a_n\vec{e}_n$, entonces

$$\phi_i(\vec{u}) = a_1\phi_i(\vec{e}_1) + \cdots + a_i\phi_i(\vec{e}_i) + \cdots + a_n\phi_n(\vec{e}_n) = a_i$$

y por tanto

$$\vec{u} = \phi_1(\vec{u})\vec{e}_1 + \cdots + \phi_i(\vec{u})\vec{e}_i + \cdots + \phi_n(\vec{u})\vec{e}_n.$$

Entonces

$$f(\vec{u}) = \phi_1(\vec{u})f(\vec{e}_1) + \cdots + \phi_i(\vec{u})f(\vec{e}_i) + \cdots + \phi_n(\vec{u})f(\vec{e}_n);$$

es decir,

$$f = f(\vec{e}_1)\phi_1 + \cdots + f(\vec{e}_i)\phi_i + \cdots + f(\vec{e}_n)\phi_n.$$

Luego $\dim(\mathbf{E}^*) = n$. ✔

31 Construir una base para el espacio dual de $\mathbb{R}^2$, $(\mathbb{R}^2)^*$, a partir de la base $\{(1,2),(1,1)\}$ de $\mathbb{R}^2$ por medio del procedimiento dado en el ejercicio precedente; es decir, encontrar la base dual de la base $\{(1,2),(1,1)\}$.

Solución Por el procedimiento del ejercicio precedente $\phi_1(1,2) = 1$, $\phi_1(1,1) = 0$, $\phi_2(1,2) = 0$ y $\phi_2(1,1) = 1$. Puesto que los ϕ_i son lineales se debe tener

$$\begin{aligned} \phi_1(x,y) &= ax + by, \\ \phi_2(x,y) &= cx + dy \end{aligned}$$

para ciertas constantes a, b, c, d. Entonces

$$\begin{aligned}1 &= \phi_1(1,2) = a+2b,\\ 0 &= \phi_1(1,1) = a+b,\\ 0 &= \phi_2(1,2) = a+2d,\\ 1 &= \phi_2(1,1) = c+d\end{aligned}$$

y dado que

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0\\ 1 & 1 & 0 & 1\end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0\\ 0 & -1 & -1 & 1\end{array}\right],$$

se obtiene

$$\begin{bmatrix} a\\ b\end{bmatrix} = \begin{bmatrix} -1\\ 1\end{bmatrix} \quad \text{y} \quad \begin{bmatrix} c\\ d\end{bmatrix} = \begin{bmatrix} 2\\ -1\end{bmatrix}$$

y por tanto

$$\begin{aligned}\phi_1(x,y) &= -x+y,\\ \phi_2(x,y) &= 2x-y\end{aligned}$$

son los elementos de la base del espacio dual $(\mathbb{R}^2)^*$. ✔

32 Si $\mathbf{E}$ es un espacio vectorial y $\mathbf{E}^*$ su espacio dual, se denota por $\mathbf{E}^{**}$ el espacio dual de $\mathbf{E}^*$; esto es, $\mathbf{E}^{**} = (\mathbf{E}^*)^* = \mathscr{L}(\mathbf{E}^*, \mathbb{R})$. Al espacio $\mathbf{E}^{**}$ se le llama *espacio bidual* (o *doble dual*) del espacio $\mathbf{E}$. Sea $u \in \mathbf{E}$ un elemento fijo (que se ha escrito sin la flecha encima por simplicidad en la notación); se define $\widehat{u} : \mathbf{E}^* \to \mathbb{R}$ por $\widehat{u}(f) = f(u)$.

(a) Probar que $\widehat{u} \in \mathbf{E}^{**}$.

(b) Sea $\Phi : \mathbf{E} \to \mathbf{E}^{**}$ definida por $\Phi(u) = \widehat{u}$. Probar que Φ es lineal.

(c) Si $\dim(\mathbf{E}) = n$ es finita, probar que Φ es un isomorfismo (el llamado *isomorfismo natural* o *canónico*) y por tanto $\mathbf{E} \cong \mathbf{E}^{**}$.

DEMOSTRACIÓN ■ *(a)* Si $f_1, f_2 \in \mathbf{E}^*$ y $\alpha, \beta \in \mathbb{R}$, entonces

$$\begin{aligned}\widehat{u}(\alpha f_1 + \beta f_2) &= (\alpha f_1 + \beta f_2)(u)\\ &= \alpha f_1(u) + \beta f_2(u)\\ &= \alpha\widehat{u}(f_1) + \beta\widehat{u}(f_2)\end{aligned}$$

y por ende $\widehat{u} \in \mathbf{E}^{**}$.

(b) Sean $u, v \in \mathbf{E}$, $\alpha, \beta \in \mathbb{R}$ y f cualquier elemento de $\mathbf{E}^*$, entonces

$$\begin{aligned}\widehat{\alpha u + \beta v}(f) &= f(\alpha u + \beta v)\\ &= f(\alpha u) + f(\beta v)\\ &= \alpha f(u) + \beta f(v)\\ &= \alpha\widehat{u}(f) + \beta\widehat{v}(f)\end{aligned}$$

y por tanto,

$$\widehat{\alpha u+\beta v} = \alpha\widehat{u}+\beta\widehat{v}; \quad \text{i.e.,}$$
$$\Phi(\alpha u+\beta v) = \alpha\Phi(u)+\beta\Phi(v).$$

(c) Sean $\{e_1,\ldots,\vec{e}_n\}$ una base de $\mathbf{E}$ y $\{\phi_1,\ldots,\phi_n\}$ la base dual correspondiente. Sea $\psi\in\mathbf{E}^{**}$ y $f\in\mathbf{E}^*$ un funcional cualquiera. Entonces, por el ejercicio resuelto 30,

$$f = f(\vec{e}_1)\phi_1+\cdots+f(\vec{e}_n)\phi_n;$$

luego

$$\begin{aligned}\psi(f) &= \psi(f(\vec{e}_1)\phi_1+\cdots+f(\vec{e}_n)\phi_n)\\ &= f(\vec{e}_1)\psi(\phi_1)+\cdots+f(\vec{e}_n)\psi(\phi_n)\\ &= f(\psi(\phi_1)\vec{e}_1+\cdots+\psi(\phi_n)\vec{e}_n).\end{aligned}$$

Entonces, si $u=\psi(\phi_1)\vec{e}_1+\cdots+\psi(\phi_n)\vec{e}_n\in\mathbf{E}$, se tiene

$$\begin{aligned}\psi(f) &= f(u)\\ &= \widehat{u}(f)\\ &= \Phi(u)(f)\end{aligned}$$

y por tanto

$$\Phi(u)=\psi.$$

Luego Φ es suprayectiva. Si $u\in\operatorname{Ker}(\Phi)$, entonces $\widehat{u}(f)=f(u)=0$ para todo funcional f; luego (cfr. el ejercicio resuelto 30 de esta sección)

$$\begin{aligned}u &= \phi_1(u)\vec{e}_1+\cdots+\phi_n(u)\vec{e}_n\\ &= \vec{0}_{\mathbf{E}}.\end{aligned}$$

Lo cual prueba que $\operatorname{Ker}(\Phi)=\{\vec{0}_{\mathbf{E}}\}$. Entonces Φ es un isomorfismo; por tanto, $\mathbf{E}\cong\mathbf{E}^{**}$. ✔

33 Sea $\mathbf{E}$ un espacio vectorial de dimensión finita n y $\mathbf{E}^*$ su espacio dual. Probar que si $\mathscr{B}^*=\{\phi_1,\ldots,\phi_n\}$ es una base de $\mathbf{E}^*$, entonces existe una base $\mathscr{B}=\{\vec{e}_1,\ldots,\vec{e}_n\}$ de $\mathbf{E}$ tal que $\mathscr{B}^*$ es la base dual de $\mathscr{B}$.

DEMOSTRACIÓN ■ Dado que $\mathbf{E}^{**}\cong\mathbf{E}$, $\dim(\mathbf{E}^{**})=n$. Sea $\{\Psi_1,\ldots,\Psi_n\}$ la base dual de la base $\{\phi_1,\ldots,\phi_n\}$ del espacio $\mathbf{E}^{**}$. Entonces,

$$\Psi_i(\phi_j)=\delta_{ij}=\begin{cases}1 \text{ si } i=j\\ 0 \text{ si } i\neq j\end{cases}$$

Por el ejercicio 32 existen únicos $\vec{e}_1,\ldots,\vec{e}_n\in\mathbf{E}$ tales que

$$\widehat{\vec{e}_i}=\Phi(\vec{e}_i)=\Psi_i$$

para cada $i = 1,\ldots,n$; y por tanto

$$\phi_j(\vec{e}_i) = \widehat{\vec{e}_i}(\phi_j) = \Psi_i(\phi_j) = \delta_{ij};$$

y ya que Φ es un isomorfismo y $\{\Psi_1,\ldots,\Psi_n\}$ es una base, $\{\vec{e}_1,\ldots,\vec{e}_n\}$ es base de $\mathbf{E}$. Luego $\{\phi_1,\ldots,\phi_n\}$ es la base dual de la base $\{\vec{e}_1,\ldots,\vec{e}_n\}$. ■

34 Sean $f_1, f_2 : \mathbb{R}^2 \to \mathbb{R}$ los funcionales lineales definidos por $f_1(x,y) = 2x - y$ y $f_2(x,y) = x + y$.

(a) Probar que $\{f_1, f_2\}$ es una base de $(\mathbb{R}^2)^*$.

(b) Encontrar una base $\{\vec{u}_1, \vec{u}_2\}$ de $\mathbb{R}^2$ cuya base dual sea $\{f_1, f_2\}$.

Solución *(a)* Ya que $\dim(\mathbb{R}^2)^* = \dim(\mathbb{R}^2) = 2$, basta probar que f_1y f_2 son L.I. Sean $\lambda_1, \lambda_2 \in \mathbb{R}$ tales que

$$\lambda_1 f_1 + \lambda_2 f_2 = \theta$$

es decir,

$$\lambda_1 f_1(x,y) + \lambda_2 f_2(x,y) = 0 \quad \forall (x,y) \in \mathbb{R}^2.$$

Entonces, en particular,

$$\begin{aligned} 0 &= \lambda_1 f_1(1,0) + \lambda_2 f_2(1,0) = 2\lambda_1 + \lambda_2 \quad \text{y} \\ 0 &= \lambda_1 f_1(0,1) + \lambda_2 f_2(0,1) = -\lambda_1 + \lambda_2\,; \end{aligned}$$

es decir,

$$\begin{aligned} 2\lambda_1 + \lambda_2 &= 0 \\ -\lambda_1 + \lambda_2 &= 0 \end{aligned}$$

de donde $\lambda_1 = \lambda_2 = 0$ y, por tanto, f_1 y f_2 son L.I.

(b) Por el ejercicio precedente existe una base $\{\vec{e}_1, \vec{e}_2\}$, $\vec{e}_1 = (a,b)$ y $\vec{e}_2 = (c,d)$ cuya base dual es $\{f_1, f_2\}$; entonces

$$\begin{aligned} 1 &= f_1(a,b) = 2a - b, \\ 0 &= f_1(c,d) = 2c - d, \\ 0 &= f_2(a,b) = a + b, \\ 1 &= f_2(c,d) = c + d. \end{aligned}$$

Y ya que

$$\left[\begin{array}{rr|rr} 2 & -1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rr|rr} 2 & -1 & 1 & 0 \\ 0 & -3 & 1 & -2 \end{array}\right]$$

se tiene

$$\begin{aligned} (a,b) &= (1/3, -1/3) \text{ y} \\ (c,d) &= (1/3, 2/3); \end{aligned}$$

es decir,

$$\{\vec{e}_1, \vec{e}_2\} = \{(1/3, -1/3), (1/3, 2/3)\}. \quad ■$$

35 (**Hiperespacio**). Sean $\mathbf{E}$ un espacio vectorial no trivial y $H < \mathbf{E}$. Se dice que H es un *hiperespacio* de $\mathbf{E}$ si se cumplen las siguientes condiciones:

(i) $H \neq \mathbf{E}$.

(ii) Si $S < \mathbf{E}$ y $H \subset S$, entonces $S = H$ o $S = \mathbf{E}$.

Es decir, H es un subespacio propio maximal. Mostrar que un subespacio H es un hiperespacio de $\mathbf{E}$ si y sólo si existe $\vec{u}_0 \in \mathbf{E} - H$ tal que

$$\mathbf{E} = H + \text{gn}(\vec{u}_0).$$

DEMOSTRACIÓN ■ ($\Rightarrow$) Como $H \neq \mathbf{E}$, existe $\vec{u}_0 \in \mathbf{E} - H$. Entonces

$$H \subsetneq H + \text{gn}(\vec{u}_0) \subset \mathbf{E}$$

y ya que H es un hiperespacio, se deduce que $H + \text{gn}(\vec{u}_0) = \mathbf{E}$. ($\Leftarrow$) Sea $\vec{u}_0 \in \mathbf{E} - H$ tal que $\mathbf{E} = H + \text{gn}(\vec{u}_0)$, entonces $H \neq \mathbf{E}$. Sea $S < \mathbf{E}$ tal que $H \subset S$. Dado que $\mathbf{E} = H + \text{gn}(\vec{u}_0)$, todos los elementos de S tienen la forma $\vec{u} = \vec{u}_1 + \beta\vec{u}_0$ para algún $\vec{u}_1 \in H$ y para cierto $\beta \in \mathbb{R}$. Si para alguno de éstos $\beta \neq 0$, se tiene $\vec{u}_0 = (1/\beta)(\vec{u} - \vec{u}_1) \in S$ y entonces $\mathbf{E} = H + \text{gn}(\vec{u}_0) = S$. En caso contrario $\vec{u} \in H$ para todo $\vec{u} \in S$; es decir, $S = H$. ■

36 Sean $\mathbf{E}$ un espacio vectorial no trivial y $f \in \mathbf{E}^*$. Demostrar lo siguiente (cfr. ejercicio resuelto 9 de esta sección):

(a) Si f es un funcional lineal en $\mathbf{E}$ no nulo ($f \neq \theta$, el funcional lineal constante cero), entonces $H = \text{Ker}(f)$ es un hiperespacio.

(b) Si H es un hiperespacio existe un funcional lineal f en $\mathbf{E}^*$ no nulo tal que $\text{Ker}(f) = H$.

DEMOSTRACIÓN ■ *(a)* Sea $\vec{u}_0 \in \mathbf{E}$ tal que $f(\vec{u}_0) \neq 0$. Si $\vec{u} \in \mathbf{E}$, entonces

$$\vec{u} - \frac{f(\vec{u})}{f(\vec{u}_0)}\vec{u}_0 \in H = \text{Ker}(f)$$

pues

$$f(\vec{u} - \frac{f(\vec{u})}{f(\vec{u}_0)}\vec{u}_0) = f(\vec{u}) - \frac{f(\vec{u})}{f(\vec{u}_0)}f(\vec{u}_0) = 0.$$

Por tanto, existe $\vec{u}_1 \in H$ tal que

$$\vec{u} - \frac{f(\vec{u})}{f(\vec{u}_0)}\vec{u}_0 = \vec{u}_1;$$

luego

$$\vec{u} = \vec{u}_1 + \frac{f(\vec{u})}{f(\vec{u}_0)}\vec{u}_0 \in H + \text{gn}(\vec{u}_0).$$

Por el ejercicio precedente se concluye que $H = \text{Ker}(f)$ es un hiperespacio.

(*b*) Sea H un hiperespacio y $\vec{u}_0 \in \mathbf{E} - H$. Entonces, por el ejercicio precedente,

$$\mathbf{E} = H + \text{gn}(\vec{u}_0).$$

y más aún, puesto que $H \cap \text{gn}(\vec{u}_0) = \{\vec{0}_{\mathbf{E}}\}$, se tiene

$$\mathbf{E} = H \oplus \text{gn}(\vec{u}_0).$$

Entonces para todo $\vec{u} \in \mathbf{E}$ existen únicos $\vec{u}_1 \in H$ y $\beta \in \mathbb{R}$ tales que

$$\vec{u} = \vec{u}_1 + \beta\vec{u}_0.$$

Sea $f : \mathbf{E} \to \mathbb{R}$ definida por

$$f(\vec{u}) = \beta.$$

Si $\vec{u}_1 + \beta_1 u_0, \vec{u}_2 + \beta_2\vec{u}_0 \in \mathbf{E} = H \oplus \text{gn}(\vec{u}_0)$, y $\lambda \in \mathbb{R}$, entonces

$$\begin{aligned} f(\vec{u}_1 + \beta_1 u_0 + \vec{u}_2 + \beta_2\vec{u}_0) &= f(\vec{u}_1 + \vec{u}_2 + (\beta_1 + \beta_2)\vec{u}_0) \\ &= \beta_1 + \beta_2 \\ &= f(\vec{u}_1 + \beta_1 u_0) + f(\vec{u}_2 + \beta_2\vec{u}_0) \end{aligned}$$

y

$$\begin{aligned} f(\lambda(\vec{u}_1 + \beta_1 u_0)) &= f(\lambda\vec{u}_1 + \lambda\beta_1 u_0) \\ &= \lambda\beta_1 \\ &= \lambda f(\vec{u}_1 + \beta_1 u_0). \end{aligned}$$

Finalmente,

$$\begin{aligned} &\vec{u} = \vec{u}_1 + \beta\vec{u}_0 \in \text{Ker}(f) \\ &\Leftrightarrow \beta = 0 \\ &\Leftrightarrow \vec{u} = \vec{u}_1 \in H. \end{aligned}$$

Es decir,

$$\text{Ker}(f) = H. \quad \blacksquare$$

37 (**Teorema de representación de Riesz**). Sean $\mathbf{E}$ un espacio con producto interior $\langle\cdot,\cdot\rangle$ y $f \in \mathbf{E}^*$ un funcional lineal. Si $\mathbf{E}$ tiene dimensión finita n, probar que existe un único $\vec{u}_f \in \mathbf{E}$ tal que $f(\vec{x}) = \langle\vec{x}, \vec{u}_f\rangle$ para todo $\vec{x} \in \mathbf{E}$.

DEMOSTRACIÓN ■ Sea $\{\vec{e}_1, \ldots, \vec{e}_n\}$ una base ortonormal de $\mathbf{E}$ y $\vec{x} \in \mathbf{E}$, entonces

$$\vec{x} = \langle\vec{x}, \vec{e}_1\rangle\,\vec{e}_1 + \cdots + \langle\vec{x}, \vec{e}_n\rangle\,\vec{e}_n$$

y por tanto,

$$\begin{aligned} f(\vec{x}) &= f(\langle\vec{x},\vec{e}_1\rangle\vec{e}_1+\cdots+\langle\vec{x},\vec{e}_n\rangle\vec{e}_n) \\ &= \langle\vec{x},\vec{e}_1\rangle f(\vec{e}_1)+\cdots+\langle\vec{x},\vec{e}_n\rangle f(\vec{e}_n) \\ &= \langle\vec{x},f(\vec{e}_1)\vec{e}_1\rangle+\cdots+\langle\vec{x},f(\vec{e}_n)\vec{e}_n\rangle \\ &= \langle\vec{x},f(\vec{e}_1)\vec{e}_1+\cdots+f(\vec{e}_n)\vec{e}_n\rangle \end{aligned}$$

Luego, si $\vec{u}_f = f(\vec{e}_1)\vec{e}_1+\cdots+f(\vec{e}_n)\vec{e}_n$, se tiene

$$f(\vec{x}) = \langle\vec{x},\vec{u}_f\rangle$$

para todo $\vec{x}\in\mathbf{E}$. Sea $\vec{v}\in\mathbf{E}$ tal que

$$\langle\vec{x},\vec{u}_f\rangle = f(\vec{x}) = \langle\vec{x},\vec{v}\rangle$$

para todo $\vec{x}\in\mathbf{E}$, entonces

$$\langle\vec{x},\vec{u}_f-\vec{v}\rangle = 0 \quad \forall\vec{x}\in\mathbf{E}$$

y por ende

$$\vec{u}_f = \vec{v}. \quad \blacksquare$$

38 (**Operador adjunto**). Sea $\mathbf{E}$ un espacio vectorial con producto interior $\langle\cdot,\cdot\rangle$. Si $\mathbf{E}$ tiene dimensión finita y $T:\mathbf{E}\to\mathbf{E}$ es un operador lineal, demostrar que existe un único operador lineal $T^*:\mathbf{E}\to\mathbf{E}$ tal que

$$\langle T(\vec{u}),\vec{v}\rangle = \langle\vec{u},T^*(\vec{v})\rangle$$

para todo $\vec{u},\vec{v}\in\mathbf{E}$. Al operador T^* se le llama el operador *adjunto* de T.

DEMOSTRACIÓN ■ Es fácil probar que para cada $\vec{v}\in\mathbf{E}$ la aplicación $f(\vec{u}) = \langle T(\vec{u}),\vec{v}\rangle$ es lineal y por tanto un elemento de $\mathbf{E}^*$. Por el ejercicio resuelto 37 (teorema de representación de Riesz), existe, para cada $\vec{v}\in\mathbf{E}$, un único $\vec{v}^*\in\mathbf{E}$ tal que

$$f(\vec{u}) = \langle\vec{u},\vec{v}^*\rangle \quad \forall\vec{u}\in\mathbf{E}.$$

Sea $T^*:\mathbf{E}\to\mathbf{E}$ definida por $T(\vec{v}) = v^*$. Entonces

$$\langle T(\vec{u}),\vec{v}\rangle = f(\vec{u}) = \langle\vec{u},\vec{v}^*\rangle = \langle\vec{u},T^*(\vec{v})\rangle$$

para todo $\vec{u},\vec{v}\in\mathbf{E}$. Si $\vec{v}_1,\vec{v}_2\in\mathbf{E}$, $\alpha,\beta\in\mathbb{R}$ y $\vec{u}$ es cualquier vector en $\mathbf{E}$, se tiene

$$\begin{aligned} \langle\vec{u},T^*(\alpha\vec{v}_1+\beta\vec{v}_2)\rangle &= \langle T(\vec{u}),\alpha\vec{v}_1+\beta\vec{v}_2\rangle \\ &= \alpha\langle T(\vec{u}),\vec{v}_1\rangle+\beta\langle T(\vec{u}),\vec{v}_2\rangle \\ &= \alpha\langle\vec{u},T^*(\vec{v}_1)\rangle+\beta\langle\vec{u},T^*(\vec{v}_2)\rangle \\ &= \langle\vec{u},\alpha T^*(\vec{v}_1)\rangle+\langle\vec{u},\beta T^*(\vec{v}_2)\rangle \\ &= \langle\vec{u},\alpha T^*(\vec{v}_1)+\beta T^*(\vec{v}_2)\rangle \end{aligned}$$

de donde

$$\langle \vec{u}, T^*(\alpha\vec{v}_1 + \beta\vec{v}_2) - (\alpha T^*(\vec{v}_1) + \beta T^*(\vec{v}_2)) \rangle = 0 \quad \forall \vec{u} \in \mathbf{E};$$

luego

$$T^*(\alpha\vec{v}_1 + \beta\vec{v}_2) - (\alpha T^*(\vec{v}_1) + \beta T^*(\vec{v}_2)) = \vec{0}_{\mathbf{E}}$$

y por tanto

$$T^*(\alpha\vec{v}_1 + \beta\vec{v}_2) = \alpha T^*(\vec{v}_1) + \beta T^*(\vec{v}_2).$$

Lo cual prueba que T^* es lineal. Si T_1 es un operador lineal tal que para todo $\vec{u}, \vec{v} \in \mathbf{E}$

$$\langle T(\vec{u}), \vec{v} \rangle = \langle \vec{u}, T_1(\vec{v}) \rangle$$

entonces

$$\langle \vec{u}, T_1(\vec{v}) \rangle = \langle \vec{u}, T^*(\vec{v}) \rangle$$

por lo que

$$\langle \vec{u}, T_1(\vec{v}) - T^*(\vec{v}) \rangle = 0 \quad \forall \vec{u}, \vec{v} \in \mathbf{E};$$

de donde se desprende

$$T_1 = T^*. \quad ■$$

39 Sean $\mathbf{E}$, T y T^* como en el ejercicio anterior. Probar que si $\mathscr{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ es una base ortonormal de $\mathbf{E}$, A es la representación matricial de T relativa a esta base y B es la representación matricial de T^* relativa a la misma base, entonces $B = A^t$.

DEMOSTRACIÓN ■ Por el ejercicio resuelto 21 de este capítulo, la representación matricial de T está dada por la matriz $A = [a_{ij}]^t$ donde $a_{ij} = \langle T(\vec{e}_i), \vec{e}_j \rangle$; por el mismo ejercicio $B = [T^*]_{\mathscr{B}} = [b_{ij}]^t$ donde

$$\begin{aligned} b_{ij} &= \langle T^*(\vec{e}_i), \vec{e}_j \rangle \\ &= \langle \vec{e}_i, T(\vec{e}_j) \rangle \\ &= a_{ji} \end{aligned}$$

y por tanto

$$\begin{aligned} B &= [b_{ij}]^t \\ &= A^t. \quad ■ \end{aligned}$$

40 (**Anuladores**). Si $\mathbf{E}$ es un espacio vectorial y S es un subconjunto no vacío de $\mathbf{E}$, se define

$$S^0 = \{\phi \in \mathbf{E}^* \mid \phi(\vec{u}) = 0 \quad \forall \vec{u} \in S\}.$$

(a) Probar que S^0 es un subespacio de $\mathbf{E}^*$. A S^0 se le llama el anulador de S.

(b) Si $\mathbf{E}$ tiene dimensión finita y S es un subespacio de $\mathbf{E}$, demostrar que

$$\dim(S) + \dim(S^0) = \dim(\mathbf{E}).$$

DEMOSTRACIÓN ■ *(a)* Claramente el operador constante cero, θ, pertenece a S^0. Si $f, g \in S^0$ y $\alpha, \beta \in \mathbb{R}$, entonces, para todo $\vec{u} \in S$,

$$\begin{aligned}(\alpha f + \beta g)(\vec{u}) &= \alpha f(\vec{u}) + \beta g(\vec{u})\\ &= 0 + 0\\ &= 0.\end{aligned}$$

(b) Sea $\{\vec{u}_1, \ldots, \vec{u}_n\}$ una base de S; se completa ésta a una base

$$\{\vec{u}_1, \ldots, \vec{u}_m, \vec{v}_{m+1}, \ldots, \vec{v}_n\}$$

de $\mathbf{E}$, y sea $\{\phi_1, \ldots, \phi_n\}$ su base dual. Entonces si $k \geq m+1$

$$\phi_k(\vec{u}_j) = 0 \, \forall j$$

y por tanto $\phi_{m+1}, \ldots, \phi_n \in S^0$. Si $\vec{u} \in \mathbf{E}$ y $\phi \in S^0$, entonces

$$\begin{aligned}\phi(\vec{u}) &= \phi(\vec{u}_1)\phi_1(\vec{u}) + \cdots + \phi(\vec{u}_m)\phi_m(\vec{u}) +\\ &\quad \phi(\vec{v}_{m+1})\phi_{m+1}(\vec{u}) + \cdots + \phi(\vec{v}_n)\phi_m(\vec{u})\\ &= 0 + \cdots + 0 + \phi(\vec{v}_{m+1})\phi_{m+1}(\vec{u}) + \cdots + \phi(\vec{v}_n)\phi_m(\vec{u})\\ &= \phi(\vec{v}_{m+1})\phi_{m+1}(\vec{u}) + \cdots + \phi(\vec{v}_n)\phi_m(\vec{u}).\end{aligned}$$

Luego $\phi_{m+1}, \ldots, \phi_n$ generan a S^0 y ya que las funciones $\phi_{m+1}, \ldots, \phi_n$ son L.I., $\{\phi_{m+1}, \ldots, \phi_n\}$ es una base de S^0. De donde

$$\dim(\mathbf{E}) - \dim(S) = \dim(S^0). \quad ■$$

41 (**Primer teorema de isomorfismo**). Sean $\mathbf{E}$ y $\mathbf{F}$ un par de espacios vectoriales y $T : \mathbf{E} \to \mathbf{F}$ una transformación lineal suprayectiva. Si $H = \operatorname{Ker}(T)$, probar que

$$\mathbf{E}/H \cong \mathbf{F}$$

donde $\mathbf{E}/H$ es el espacio cociente (cfr. el ejercicio resuelto 29 del capítulo 3).

DEMOSTRACIÓN ■ Sea $\Phi : \mathbf{E}/H \to \mathbf{F}$ la transformación definida, para cada $[\vec{u}] \in \mathbf{E}/H$, por

$$\Phi([\vec{u}]) = T(\vec{u});$$

entonces,

(a) Φ está bien definida: si $[\vec{u}] \in \mathbf{E}/H$ y $\vec{v} \in [\vec{u}]$, entonces $\vec{u} - \vec{v} \in H = \operatorname{Ker}(T)$ y por tanto $T(\vec{u} - \vec{v}) = \vec{0}_{\mathbf{F}}$; i.e., $T(\vec{u}) = T(\vec{v})$.

(b) Φ es lineal: si $[\vec{u}], [\vec{v}] \in \mathbf{E}/H$ y $\alpha, \beta \in \mathbb{R}$, se tiene

$$\begin{aligned}\Phi(\alpha[\vec{u}] + \beta[\vec{v}]) &= \Phi([\alpha\vec{u} + \beta\vec{v}])\\ &= T(\alpha\vec{u} + \beta\vec{v})\\ &= \alpha T(\vec{u}) + \beta T(\vec{v})\\ &= \alpha\Phi([\vec{u}]) + \beta\Phi([\vec{v}]).\end{aligned}$$

(c) Φ es inyectiva: si $[\vec{u}] \in \text{Ker}(\Phi)$, entonces

$$\begin{aligned} T(\vec{u}) &= \Phi([\vec{u}]) \\ &= \vec{0}_{\mathbf{F}} \end{aligned}$$

y por tanto $\vec{u} \in \text{Ker}(T) = H$; así que $[\vec{u}] = H = [\vec{0}_{\mathbf{E}}] = \vec{0}_{\mathbf{E}/H}$.

(d) Φ es suprayectiva: si $\vec{w} \in \mathbf{F}$, dado que T es suprayectiva, existe $\vec{u} \in \mathbf{E}$ tal que $T(\vec{u}) = \vec{w}$, luego

$$\Phi([\vec{u}]) = T(\vec{u}) = \vec{w}. \quad ■$$

42 Sean $\mathbf{P}$ el espacio de polinomios y $p(x)$ un polinomio dado de grado $\text{grad}(p) = n$.

(a) Probar que $H = \{pf \mid f \in \mathbf{P}\}$ es un subespacio de $\mathbf{P}$.

(b) Mostrar que $\mathbf{P}/H \cong \mathbb{R}^n$.

DEMOSTRACIÓN ■ *(a)* Si θ es el polinomio constante cero, entonces $p\theta = \theta$ y por tanto $\theta \in H$. Sean $pf_1, pf_2 \in H$ y $\alpha, \beta \in \mathbb{R}$, entonces

$$\alpha(pf_1) + \beta(pf_2) = p(\alpha f_1 + \beta f_2) \in H.$$

(b) Por el algoritmo de división para cada $g(x) \in \mathbf{P}$ existen únicos polinomios $q(x)$ y $r(x)$, $r(x) = 0$ o $0 \leq \text{grad}(r) \leq n-1$ ($\text{grad}(g)$ representa el grado del polinomio g), tales que $g(x) = p(x)q(x) + r(x)$. Sea $T : \mathbf{P} \to \mathbf{P}_{n-1}$ dada por

$$T(g(x)) = r(x).$$

Si $g_1(x) = p(x)q_1(x) + r_1(x)$ y $g_2(x) = p(x)q_2(x) + r_2(x)$ y $\alpha \in \mathbb{R}$, con $0 \leq \text{grad}(r_1), \text{grad}(r_2) \leq n-1$ o $r_1(x) = 0$ y/o $r_2(x) = 0$, entonces

$$\begin{aligned} g_1(x) + g_2(x) &= p(x)(q_1(x) + q_2(x)) + r_1(x) + r_2(x), \\ \alpha g_1(x) &= p(x)(\alpha q(x)) + (\alpha r_1(x)), \\ 0 &\leq \text{grad}(r_1 + r_2) \leq n-1 \ (\text{o } r_1 + r_2 = \theta) \quad \text{y} \\ 0 &\leq \text{grad}(\alpha r_1) \leq n-1 \end{aligned}$$

por lo que

$$\begin{aligned} T(g_1(x) + g_2(x)) &= r_1(x) + r_2(x) \\ &= T(g_1(x)) + T(g_2(x)) \quad \text{y} \\ T(\alpha g_1(x)) &= \alpha r_1(x) \\ &= \alpha T(g_1(x)). \end{aligned}$$

Lo cual prueba que T es lineal. Si $r(x) \in \mathbf{P}_{n-1}$, entonces $r(x) = p(x)\theta(x) + r(x)$ y por tanto $T(r(x)) = r(x)$; así que T es suprayectiva. Por otra parte

$$\begin{aligned} g(x) \in \operatorname{Ker}(T) &\Leftrightarrow T(g(x)) = r(x) = 0 \\ &\Leftrightarrow g(x) = p(x)q(x) \\ &\Leftrightarrow g(x) \in H. \end{aligned}$$

Por el ejercicio anterior

$$\mathbf{P}/H \cong \mathbf{P}_{n-1}$$

y dado que $\dim(\mathbf{P}_{n-1}) = n = \dim(\mathbb{R}^n)$, del teorema 5.24 (cfr. pág. 456) se tiene $\mathbf{P}_{n-1} \cong \mathbb{R}^n$ y entonces

$$\mathbf{P}/H \cong \mathbb{R}^n. \quad ■$$

Valores y vectores propios, diagonalización

43 Sean $\mathbf{E}$ un espacio vectorial y S un subespacio de él. Si $T : S \to \mathbf{E}$ es una transformación lineal también se acostumbra decir que T es un operador lineal en $\mathbf{E}$ (aunque S sea distinto de $\mathbf{E}$). Un escalar λ es valor propio de T si existe $\vec{u} \in S - \{\vec{0}_{\mathbf{E}}\}$ tal que $T(\vec{u}) = \lambda\vec{u}$. Sean $\mathbf{E} = C(-\infty,\infty)$ el espacio vectorial de funciones continuas en todo punto $x \in \mathbb{R}$ y S el conjunto de funciones $f \in \mathbf{E}$ tales que $\int_{-\infty}^{x} f(t)dt$ converge para todo $x \in \mathbb{R}$.

(a) Mostrar que S es un subespacio de $\mathbf{E}$.

(b) Se define $T : S \to \mathbf{E}$, para cada $f \in S$, $T(f) = g$ donde

$$g(x) = \int_{-\infty}^{x} f(t)dt.$$

Mostrar T es un operador lineal de S en $\mathbf{E}$.

(c) Encontrar los valores propios y vectores propios correspondientes.

Solución *(a)* Claramente la función constante cero, θ, pertenece a S. Sean $f_1, f_2 \in S$; $\alpha, \beta \in \mathbb{R}$ un par de escalares y $x \in \mathbb{R}$, entonces

$$\begin{aligned} \int_{-\infty}^{x} (\alpha f_1 + \beta f_2)(t)dt &= \lim_{r\to\infty} \int_{-r}^{x} (\alpha f_1(t) + \beta f_2(t))dt \\ &= \alpha \lim_{r\to\infty} \int_{-r}^{x} f_1(t)dt + \beta \lim_{r\to\infty} \int_{-r}^{x} f_2(t)dt \\ &= \alpha \int_{-\infty}^{x} f_1(t)dt + \beta \int_{-\infty}^{x} f_2(t)dt \end{aligned}$$

y por tanto $\int_{-\infty}^{x} (\alpha f_1 + \beta f_2)t dt$ converge para todo $x \in \mathbb{R}$; luego $\alpha f_1 + \beta f_2 \in S$; lo cual prueba que S es subespacio de $\mathbf{E}$.

(b) Si $f \in S$, entonces f es continua en todo punto y $\int_{-\infty}^{x} f(t)dt$ converge para todo $x \in \mathbb{R}$. Dado que f es continua en todo punto de $\mathbb{R}$, la función $G(x) = \int_{0}^{x} f(t)dt$ es continua en todo $x \in \mathbb{R}$. Si $\alpha = \int_{-\infty}^{0} f(t)dt$, entonces

$$g(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} f(t)dt + \int_{0}^{x} f(t)dt$$
$$= \alpha + G(x)$$

y por tanto $g \in \mathbf{E} = C(-\infty,\infty)$. Lo cual prueba que efectivamente $T(f) \in \mathbf{E} = C(-\infty,\infty)$ para toda $f \in S$. Sean $a,b \in \mathbb{R}$ y $f_1, f_2 \in S$, entonces, para todo $x \in \mathbb{R}$,

$$T(af_1 + bf_2)(x) = \int_{-\infty}^{x} (af_1 + bf_2)(t)dt$$
$$= a\int_{-\infty}^{x} f_1(t)dt + b\int_{-\infty}^{x} f_2(t)dt$$
$$= aT(f_1)(x) + bT(f_2)(x);$$

esto es,

$$T(af_1 + bf_2) = aT(f_1) + bT(f_2).$$

(c) Si $\int_{-\infty}^{x} f(t)dt = 0$ para todo $x \in \mathbb{R}$ y $a,b \in \mathbb{R}$ son cualquier par de números reales, con $a < b$, entonces

$$\int_{a}^{b} f(t)dt = \int_{-\infty}^{b} f(t)dt - \int_{-\infty}^{a} f(t)dt = 0.$$

Así que $\int_{a}^{b} f(t)dt = 0$ en todo intervalo $[a,b]$. Si existe $x_0 \in \mathbb{R}$ tal que $f(x_0) \neq 0$, digamos $f(x_0) > 0$, entonces, por continuidad de f, existe un intervalo $[x_0-\delta, x_0+\delta]$ tal que $f(x) > 0$ en todo este intervalo, luego $\int_{x_0-\delta}^{x_0+\delta} f(t)dt > 0$. Lo cual es una contradicción al hecho descubierto de que la integral de f es nula en todo intervalo cerrado. Por tanto, $\int_{-\infty}^{x} f(t)dt = 0$ para todo $x \in \mathbb{R}$ implica $f(x) = 0$ para todo punto x. Esto significa que $\lambda = 0$ no puede ser valor propio de T; pues la única función en S tal que $T(f) = 0$, es la función $f(x) = 0$ para todo $x \in \mathbb{R}$. Si $G(x) = \int_{0}^{x} f(t)dt$ por el teorema fundamental del cálculo, ya que f es continua en todo punto, G es derivable y además $G'(x) = f(x)$ para todo $x \in \mathbb{R}$. Por el inciso anterior, si $g = T(f)$,

$$g(x) = \alpha + G(x)$$

donde $\alpha = \int_{-\infty}^{0} f(t)dt$; luego g es derivable en todo punto y

$$g'(x) = G'(x) = f(x)$$

para todo $x \in \mathbb{R}$. Si $\lambda \neq 0$ satisface $T(f) = \lambda f$ se debe tener

$$\int_{-\infty}^{x} f(t)dt = \lambda f(x)$$

para todo punto x. Al derivar respecto a x en ambos lados de la precedente igualdad se obtiene

$$f(x) = \lambda f'(x)$$

esto es,

$$\frac{df}{dx} = \frac{1}{\lambda} f$$

que, al separar variables (suponiendo $f \neq 0$), produce

$$\int \frac{df}{f} = \frac{1}{\lambda}\int dx;$$

es decir,

$$\ln|f| = \frac{1}{\lambda}x + C$$

y por tanto

$$f = Ae^{x/\lambda}$$

para alguna constante $A \neq 0$. Si $\lambda > 0$,

$$\begin{aligned} g(x) &= \int_{-\infty}^{x} Ae^{t/\lambda}dt \\ &= A\lim_{r\to\infty}\int_{-r}^{x} e^{t/\lambda}dt \\ &= A\lambda\lim_{r\to\infty}\left[e^{x/\lambda} - e^{-r/\lambda}\right] \\ &= A\lambda e^{x/\lambda} \\ &= \lambda f(x) \end{aligned}$$

y $\int_{-\infty}^{x} Ae^{t/\lambda}dt$ diverge si $\lambda < 0$. Así, todo $\lambda > 0$ es valor propio de T con vectores propios correspondientes $f(x) = Ae^{x/\lambda}$ para cualquier constante A distinta de cero (cfr. ejemplo 5.44, pág. 460). ✔

44 Sean $T, R : \mathbf{E} \to \mathbf{E}$ dos operadores lineales suprayectivos en el espacio $\mathbf{E}$. Demostrar que TR y RT tienen los mismos valores propios, donde $TR = T \circ R$ y $RT = R \circ T$ (cfr. el ejercicio resuelto 12 de este capítulo).

DEMOSTRACIÓN ■ Sea λ un valor propio de TR, se prueba entonces que λ también es valor propio de RT. La demostración de que todo valor propio de RT es también valor propio de TR es completamente simétrica.

(a) Si $\lambda = 0$, entonces el operador TR no es inyectivo ($\operatorname{Ker}(TR - 0I) \neq \{\vec{0}_{\mathbf{E}}\}$, cfr. nota 5, pág. 459); y por tanto, del ejercicio resuelto 13*(b)* de este apartado, al menos uno de los dos operadores T, R no es inyectivo; entonces, del ejercicio resuelto 13*(c)*, RT no es inyectivo, luego $\lambda = 0$ es valor propio de RT pues al no ser inyectivo $\operatorname{Ker}(TR - 0I) \neq \{\vec{0}_{\mathbf{E}}\}$.

(b) Suponga $\lambda \neq 0$ y sea $\vec{u} \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$ tal que $(TR)(\vec{u}) = \lambda\vec{u}$, entonces

$$T(R(\vec{u})) = \lambda\vec{u}.$$

Sea $\vec{w} = R(\vec{u})$; si $\vec{w} = \vec{0}_{\mathbf{E}}$ se tiene

$$\begin{aligned} \vec{0}_{\mathbf{E}} &= T(\vec{w}) \\ &= T(R(\vec{u})) \\ &= (TR)(\vec{u}) \\ &= \lambda\vec{u} \end{aligned}$$

lo cual es imposible pues $\lambda \neq 0$ y $\vec{u} \neq \vec{0}_{\mathbf{E}}$; por tanto $\vec{w} \neq \vec{0}_{\mathbf{E}}$. Y puesto que

$$\begin{aligned}(RT)(\vec{w}) &= RT(R(\vec{u}))\\ &= R((TR)(\vec{u}))\\ &= R(\lambda\vec{u})\\ &= \lambda R(\vec{u})\\ &= \lambda\vec{w}\end{aligned}$$

se concluye que λ es también valor propio de RT. ■

45 Sea $T : \mathbf{P}_n \to \mathbf{P}_n$ el operador definido, para cada $p \in \mathbf{P}_2$, por $T(p) = q$ donde

$$q(x) = p(x) + (x+1)p'(x).$$

(a) Probar que T es un operador lineal.

(b) Encontrar los valores propios de T.

(c) Determinar si T es diagonalizable y en caso positivo hallar una diagonalización para T.

(d) ¿Existe una base $\mathscr{B}$ de $\mathbf{P}_n$ tal que la representación matricial relativa a ésta del operador T es diagonal? Si la respuesta es positiva exhibir dicha base $\mathscr{B}$.

Solución *(a)* Sean $p_1, p_2 \in \mathbf{P}_n$ y $\alpha, \beta \in \mathbb{R}$. Entonces, para todo $x \in \mathbb{R}$,

$$\begin{aligned}T(\alpha p_1 + \beta p_2)(x) &= (\alpha p_1 + \beta p_2)(x) + (x+1)(\alpha p_1 + \beta p_2')'(x)\\ &= (\alpha p_1(x) + \beta p_2(x)) + (x+1)(\alpha p_1'(x) + \beta p_2'(x))\\ &= \alpha p_1(x) + \alpha(x+1)p_1'(x) + \beta p_2(x) + \beta(x+1)p_2'(x)\\ &= \alpha T(p_1)(x) + \beta T(p_2)(x)\end{aligned}$$

y por tanto,

$$T(\alpha p_1 + \beta p_2) = \alpha T(p_1) + \beta T(p_2).$$

(b) Si $p \neq \theta$, donde θ es el polinomio constante cero, es un polinomio propio de T (un vector propio) y λ es un valor propio correspondiente, entonces

$$p + (x+1)p' = \lambda p$$

de donde

$$(x+1)\frac{dp}{dx} = (\lambda - 1)p.$$

Al separar variables se obtiene

$$\int \frac{dp}{p} = (\lambda - 1)\int \frac{dx}{x+1};$$

y por ende,

$$\ln|p| = \ln|x+1|^{\lambda-1} + C$$

para alguna constante C. Luego

$$p = C_1(x+1)^{\lambda-1}$$

con C_1 una constante distinta a cero. Dado que p debe pertenecer a $\mathbf{P}_n$, se deduce que los valores propios de T son $\lambda_k = k$ con $k = 1,2,\ldots,n+1$ y los vectores propios correspondientes a cada valor propio λ_k son los polinomios

$$p_k(x) = C(x+1)^{k-1},\ C \neq 0.$$

(c) Ya que los valores propios de T, $\lambda_k = k$, $k = 1,\ldots,n+1$, son distintos entre sí y $\dim(P_n) = n+1$, T es diagonalizable. Una diagonalización para T consiste en el par D, P donde

$$D = \operatorname{diag}(1,2,\ldots,n+1)$$

y

$$P = \begin{bmatrix} C_1 & \cdots & C_{n+1} \end{bmatrix}$$

donde la columna C_k es el vector de coordenadas del polinomio propio $p_k(x) = (1+x)^{k-1}$ correspondiente al valor propio $\lambda_k = k$; i.e., C_k es el vector que tiene como componentes los coeficientes de $(1+x)^{k-1}$:

$$C_1 = (\underbrace{1,0,\ldots,0}_{n+1})$$
$$C_k = (1, \binom{k}{1},\ldots,\binom{k}{k-2},k,1,\underbrace{0,\ldots,0}_{n-k}).$$

(d) Sí, por el inciso anterior la base es

$$\mathscr{B} = \{1, 1+x, (1+x)^2, \ldots, (1+x)^n\}.$$ ✔

Del ejercicio 46 al último ejercicio de este apartado de problemas resueltos, los espacios vectoriales, operadores lineales y matrices tratados en ellos se consideran sobre el campo de los números complejos $\mathbb{C}$ en forma implícita (cfr. 5.3.3). Sin embargo, se puede restringir a los números reales suponiendo que todos los valores propios de cualquier operador o matriz, así como todas las raíces de los polinomios involucrados son números reales.

46 Sea la matriz de tamaño $k \times k$

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & \vdots & -a_2 \\ \vdots & \vdots & \ddots & 0 & \vdots \\ 0 & 0 & 0 & 1 & -a_{k-1} \end{bmatrix}.$$

Demostrar que el polinomio característico de A es

$$p_A(\lambda) = (-1)^k(a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k). \tag{5.34}$$

DEMOSTRACIÓN ■ Se procede por inducción sobre k: si $k = 1$, entonces

$$\begin{aligned} p_A(\lambda) &= \det(-a_0 - \lambda) \\ &= -a_0 - \lambda \\ &= (-1)^1(a_0 + \lambda^1) \end{aligned}$$

y el resultado es cierto para este caso. Sea n un entero mayor a 1 y suponga que la fórmula 5.34 es válida para el caso $k = n - 1$. Entonces

$$p_A(\lambda) = \begin{vmatrix} 0-\lambda & 0 & \cdots & 0 & -a_0 \\ 1 & 0-\lambda & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & \vdots & -a_2 \\ \vdots & \vdots & \ddots & 0-\lambda & \vdots \\ 0 & 0 & 0 & 1 & -a_{n-1}-\lambda \end{vmatrix}.$$

Al desarrollar por cofactores en la primera fila y hacer uso de la hipótesis de inducción se obtiene

$$\begin{aligned} p_A(\lambda) = -\lambda &\begin{vmatrix} 0-\lambda & 0 & \cdots & 0 & -a_1 \\ 1 & 0-\lambda & \cdots & 0 & -a_2 \\ 0 & 1 & \cdots & \vdots & -a_3 \\ \vdots & \vdots & \ddots & 0-\lambda & \vdots \\ 0 & 0 & 0 & 1 & -a_{n-1}-\lambda \end{vmatrix} \\ + (-1)^{n+1}(-a_0) &\begin{vmatrix} 1 & 0-\lambda & \cdots & 0 \\ 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0-\lambda \\ 0 & 0 & 0 & 1 \end{vmatrix} \\ &= -\lambda(-1)^{n-1}(a_1 + a_2\lambda + a_3\lambda^2 + \cdots a_{n-1}\lambda^{n-2} + \lambda^{n-1}) + (-1)^n a_0 \\ &= (-1)^n(a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n). \end{aligned}$$

Luego 5.34 vale para todo $k \in \mathbb{N}$. ■

47 Sean **E** un espacio vectorial de dimensión finita, T un operador en **E** y S un subespacio T-invariante; esto es, $T(S) \subset S$. Sea $T_S : S \to S$ el operador lineal $T_S(\vec{u}) = T(\vec{u})$ para todo $\vec{u} \in S$; es decir, T_S es la restricción de T al subespacio T-invariante S. Demostrar que si $p_{T_S}(\lambda)$ es el polinomio característico de T_S y $p_T(\lambda)$ es el polinomio característico de T, entonces $p_{T_S}(\lambda)$ divide a $p_T(\lambda)$.

DEMOSTRACIÓN ■ Sea $\{\vec{u}_1, \ldots, \vec{u}_k\}$ una base de S, entonces se puede completar ésta a una base $\mathscr{B} = \{\vec{u}_1, \ldots, \vec{u}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ de **E**. Por tanto, por ser S un subespacio T-invariante se tiene

$$T(\vec{u}_j) = \sum_{i=1}^{k} a_{ij}\vec{u}_j$$

para $j = 1, \ldots, k$. Así, $A = [a_{ij}]^t$ es la representación matricial de T_S relativa a la base $\{\vec{u}_1, \ldots, \vec{u}_k\}$ y

$$[T]_{\mathscr{B}} = \begin{bmatrix} A & B \\ \mathscr{O} & C \end{bmatrix};$$

donde B es una matriz de tamaño $k\times(n-k)$, C es una matriz de tamaño $(n-k)\times(n-k)$ y $\mathscr{O}$ es la matriz cero de tamaño $(n-k)\times k$. Entonces

$$\begin{aligned} p_T(\lambda) &= \det([T]_{\mathscr{B}} - \lambda I_n) \\ &= \begin{vmatrix} A-\lambda I_k & B \\ \mathscr{O} & C-\lambda I_{n-k} \end{vmatrix} \\ &= \det(A-\lambda I_k)\det(C-\lambda I_{n-k}) \\ &= p_{T_S}(\lambda)p_C(\lambda) \end{aligned}$$

donde $p_C(\lambda)$ es el polinomio característico de la matriz C. Lo cual prueba que $p_{T_S}(\lambda)$ divide a $p_T(\lambda)$. ■

48 Sean T un operador lineal en un espacio vectorial $\mathbf{E}$ no trivial de dimensión finita y $\vec{u}\in\mathbf{E}-\{\vec{0}_{\mathbf{E}}\}$. Al subespacio S generado por los vectores $\vec{u},T(\vec{u}),T^2(\vec{u}),\ldots$; esto es $S=\mathscr{L}(\{\vec{u},T(\vec{u}),T^2(\vec{u}),\ldots\})$ (cfr. el ejercicio resuelto 31 del capítulo 3, pág. 190) se le llama el subespacio T-cíclico generado por el vector $\vec{u}$.

(a) Demostrar que existe un entero k tal que $\{\vec{u},T(\vec{u}),T^2(\vec{u}),\ldots,T^{k-1}(\vec{u})\}$ es una base de S.

(b) Probar que S es T-invariante.

(c) Sean $a_0,a_1,\ldots,a_{k-1}$ los únicos escalares tales que

$$T^k(\vec{u}) = -a_0\vec{u}-a_1T(\vec{u})-a_2T^2(u)-\cdots-a_{k-1}T^{k-1}(\vec{u}).$$

y T_S la restricción del operador T al subespacio S como en el ejercicio anterior. Demostrar que

$$p_{T_S}(\lambda) = (-1)^k(a_0+a_1\lambda+a_2\lambda^2+\cdots+a_k\lambda^k)$$

es el polinomio característico del operador T_S.

DEMOSTRACIÓN ■ *(a)* Dado que $\vec{u}\neq\vec{0}_{\mathbf{E}}$, $\{\vec{u}\}$ es L.I. y, puesto que la dimensión de $\mathbf{E}$ es finita, existe un menor entero $k\geq 1$ tal que el conjunto

$$\{\vec{u},T(\vec{u}),T^2(\vec{u}),\ldots,T^{k-1}(\vec{u})\}$$

es L.I. Entonces $\vec{u},T(\vec{u}),T^2(\vec{u}),\ldots,T^{k-1}(\vec{u}),T^k(\vec{u})$ son L.D. y por tanto

$$T^k(\vec{u})\in W=\operatorname{gn}(\vec{u},T(\vec{u}),T^2(\vec{u}),\ldots,T^{k-1}(\vec{u}))\subset S.$$

Si $\vec{v}\in W$, existen escalares $\alpha_0,\alpha_1,\ldots,\alpha_{k-1}$ tales que

$$\vec{v}=\alpha_0\vec{u}+\alpha_1T(\vec{u})+\alpha_2T^2(\vec{u})+\cdots+\alpha_{k-1}T^{k-1}(\vec{u})$$

y por tanto,

$$\begin{aligned} T(\vec{v}) &= T(\alpha_0\vec{u}+\alpha_1T(\vec{u})+\alpha_2T^2(\vec{u})+\cdots+\alpha_{k-1}T^{k-1}(\vec{u}) \\ &= \alpha_0T(\vec{u})+\alpha_1T^2(\vec{u})+\alpha_2T^3(\vec{u})+\cdots+\alpha_{k-1}T^{k-1}(\vec{u})\in W. \end{aligned}$$

Así que $W=\operatorname{gn}(\vec{u},T(\vec{u}),T^2(\vec{u}),\ldots,T^{k-1}(\vec{u}))$ es T-invariante. Dado que $\vec{u}\in W$ y este subespacio es T-invariante, se desprende que $T^r(u)\in W$ para todo $r\in\mathbb{N}$. Luego W es un subespacio que

contiene a $\{\vec{u}, T(\vec{u}), T^2(\vec{u}), \ldots\}$ y ya que $S = \mathscr{L}(\{\vec{u}, T(\vec{u}), T^2(\vec{u}), \ldots\})$ es el menor subespacio que contiene a este subconjunto (cfr. el ejercicio resuelto 31 del capítulo 3), se tiene $S \subset W$ y por tanto $S = W$. Puesto que $\mathscr{B} = \{\vec{u}, T(\vec{u}), T^2(\vec{u}), \ldots, T^{k-1}(\vec{u})\}$ es L.I. y genera a $W = S$, se concluye que es una base de S.

(b) En el inciso anterior se demostró que $S = W$ y que W es T-invariante.

(c) Sean T_S la restricción del operador T al subespacio T-cíclico S, $\mathscr{B} = \{\vec{u}, T(\vec{u}), \ldots, T^{k-1}(\vec{u})\}$ la base que se construyó en el primer inciso para éste y a_j, $j = 0, 1, \ldots, m-1$, los únicos escalares tales que $T^k(\vec{u}) = -a_0\vec{u} - a_1T(\vec{u}) - a_2T^2(\vec{u}) - \cdots - a_{k-1}T^{k-1}(\vec{u})$, entonces

$$[T_S]_{\mathscr{B}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & \vdots & -a_2 \\ \vdots & \vdots & \ddots & 0 & \vdots \\ 0 & 0 & 0 & 1 & -a_{k-1} \end{bmatrix}.$$

Por el ejercicio resuelto 46 de este apartado, el polinomio característico de esta matriz está dado por la fórmula (5.34), luego

$$p_{T_S}(\lambda) = (-1)^k(a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k). \quad ■$$

Para los ejercicios 49 a 53 considerar lo siguiente: si T es un operador lineal en un espacio vectorial de dimensión finita n y $p(x) = a_0 + a_1x + \cdots + a_mx^m$ es un polinomio, se define el operador lineal

$$p(T) = a_0I + a_1T + \cdots + a_mT^m$$

(cfr. el ejercicio propuesto 400 de este capítulo) donde I es el operador identidad en $\mathbf{E}$. Se puede interpretar a $p(T)$ como la evaluación del polinomio p en el operador T. De manera análoga si A es una matriz cuadrada de orden n, se define la matriz del mismo tamaño

$$p(A) = a_0I + a_1A + \cdots + a_mA^m,$$

donde I es la matriz identidad del mismo tamaño de A y se interpreta $p(A)$ como la evaluación del polinomio p en la matriz A.

49 Sean T un operador lineal en un espacio vectorial de dimensión finita n, A la representación matricial de T relativa a una base $\mathscr{B}$ de $\mathbf{E}$ y $p(x) = \sum_{k=0}^{m} a_kx^k$ un polinomio. Demostrar que $p(T) = \theta$, el operador constante cero, si y sólo si $p(A) = \mathscr{O}$, la matriz cero del mismo orden de A.

DEMOSTRACIÓN ■ Por el teorema 5.12 (cfr. pág. 436)

$$\begin{aligned}
[p(T)\vec{u}]_{\mathscr{B}} &= [a_0I(\vec{u}) + a_1T(\vec{u}) + a_2T^2(\vec{u}) + \cdots + a_mT^m(\vec{u})]_{\mathscr{B}} \\
&= a_0[\vec{u}]_{\mathscr{B}} + a_1[T(\vec{u})]_{\mathscr{B}} + a_2[T^2(\vec{u})]_{\mathscr{B}} + \cdots + a_m[T^m(\vec{u})]_{\mathscr{B}} \\
&= a_0[\vec{u}]_{\mathscr{B}} + a_1A[\vec{u}]_{\mathscr{B}} + a_2A^2[\vec{u}]_{\mathscr{B}} + \cdots + a_mA^m[\vec{u}]_{\mathscr{B}} \\
&= (a_0I + a_1A + a_2A^2 + \cdots + a_mA^m)[\vec{u}]_{\mathscr{B}}.
\end{aligned}$$

De donde

$$
\begin{aligned}
p(T) = \theta \Leftrightarrow p(T)\vec{u} = \vec{0}_{\mathbf{E}}\ \forall \vec{u} \in \mathbf{E} \\
&\Leftrightarrow [p(T)\vec{u}]_{\mathscr{B}} = \vec{0}_{\mathbb{R}^n}\ \forall \vec{u} \in \mathbf{E} \\
&\Leftrightarrow (a_0 I + a_1 A + a_2 A^2 + \cdots + a_m A^m)[\vec{u}]_{\mathscr{B}} = \vec{0}_{\mathbb{R}^n}\ \forall \vec{u} \in \mathbf{E} \\
&\Leftrightarrow (a_0 I + a_1 A + a_2 A^2 + \cdots + a_m A^m)\vec{x} = \vec{0}_{\mathbb{R}^n}\ \forall \vec{x} \in \mathbb{R}^n \\
&\Leftrightarrow (a_0 I + a_1 A + a_2 A^2 + \cdots + a_m A^m) = \mathscr{O} \\
&\Leftrightarrow p(A) = \mathscr{O}. \quad ■
\end{aligned}
$$

50 (**Teorema de Cayley-Hamilton**). Demostrar el siguiente teorema:

(a) Si T es un operador en un espacio vectorial de dimensión finita y $p_T(\lambda)$ es su polinomio característico, entonces

$$p_T(T) = \theta$$

donde θ es el operador constante cero; es decir,

$$p_T(T)(\vec{u}) = \vec{0}_{\mathbf{E}}\ \ \forall \vec{u} \in \mathbf{E}.$$

Lo cual significa que en un espacio vectorial de dimensión finita *todo operador es un cero de su polinomio característico*.

(b) Si A es una matriz cuadrada y $p_A(\lambda)$ es su polinomio característico, entonces

$$p_A(A) = \mathscr{O}$$

donde $\mathscr{O}$ es la matriz cero del mismo orden que la matriz A. Lo cual se interpreta diciendo que *toda matriz cuadrada es un cero de su polinomio característico*.

DEMOSTRACIÓN ■ *(a)* Sea $\vec{u} \in \mathbf{E}$, si $\vec{u} = \vec{0}_{\mathbf{E}}$, entonces claramente $p_T(T)(\vec{u}) = \vec{0}_{\mathbf{E}}$ porque T es lineal. Suponga que $\vec{u} \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$ y sea S el espacio T-cíclico generado por $\vec{u}$ considerado en el ejercicio resuelto 48 de este apartado. Por el ejercicio resuelto 48 existen un entero $k \geq 1$ y escalares únicos a_j tales que

$$\{\vec{u}, T(\vec{u}), \ldots, T^{k-1}(\vec{u})\}$$

es una base de S; S es un subespacio T-invariante;

$$
\begin{gathered}
T^k(\vec{u}) = -a_0 - a_1 T(\vec{u}) - \cdots - a_{k-1} T^{k-1}(\vec{u}); \\
p_{T_S}(\lambda) = (-1)^k (a_0 + a_1 \lambda + \cdots + a_{k-1} \lambda^{k-1} + \lambda^k)
\end{gathered}
$$

es el polinomio característico de T_S, la restricción de T al subespacio S. Entonces, por el ejercicio resuelto 47 de este capítulo, $p_{T_S}(\lambda)$ divide al polinomio característico de $p_T(\lambda)$ y por tanto existe un polinomio $q(\lambda)$ tal que

$$p_T(\lambda) = p_{T_S}(\lambda) q(\lambda).$$

Y ya que (cfr. ejercicio resuelto 48 de este apartado)

$$\begin{aligned} p_{T_S}(T)\vec{u} &= (-1)^k(a_0I + a_1T + \cdots + a_{k-1}T^{k-1} + T^k)\vec{u} \\ &= (-1)^k(a_0\vec{u} + a_1T(\vec{u}) + \cdots + a_{k-1}T^{k-1}(\vec{u}) + T^k(\vec{u})) \\ &= (-1)^k\vec{0}_{\mathbf{E}} \\ &= \vec{0}_{\mathbf{E}} \end{aligned}$$

se desprende que $p_{T_S}(T) = \theta$.

(b) Es consecuencia inmediata del inciso anterior al tomar $T = T_A$. ■

51 Sea $p(x) = a_0 + a_1x + \cdots + a_mx^m$ un polinomio de grado m.

(a) Si T es un operador lineal en un espacio vectorial $\mathbf{E}$, λ es un valor propio de T y p se anula en T; esto es, $a_0I + a_1T + \cdots a_mT^m = \theta$, el operador constante cero, demostrar que $p(\lambda) = 0$.

(b) Si A es una matriz cuadrada, λ es un valor propio de A y p se anula en A; esto es, $a_0I + a_1A + \cdots + a_mA^m = \mathscr{O}$, la matriz cero, demostrar que $p(\lambda) = 0$.

DEMOSTRACIÓN ■ *(a)* Sea $\vec{u} \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$ un vector propio correspondiente al valor propio λ. Entonces

$$\begin{aligned} \vec{0}_{\mathbf{E}} &= \theta(\vec{u}) \\ &= p(T)\vec{u} \\ &= a_0I(\vec{u}) + a_1T(\vec{u}) + a_2T^2(\vec{u}) + \cdots + a_mT^m(\vec{u}) \\ &= a_0\vec{u} + a_1\lambda\vec{u} + a_2\lambda^2\vec{u} + \cdots + a_m\lambda^m\vec{u} \\ &= (a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_m\lambda^m)\vec{u} \end{aligned}$$

y, ya que $\vec{u} \neq \vec{0}_{\mathbf{E}}$, se concluye

$$p(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_m\lambda^m = 0.$$

(b) Es inmediata del inciso anterior al poner $T = T_A$. ■

52 (**Polinomio mínimo**).

(a) Sean $\mathbf{E}$ un espacio vectorial de dimensión finita y T un operador lineal en $\mathbf{E}$.

(i) Probar que existe un polinomio mónico $\mathfrak{m}$ (el coeficiente de la mayor potencia es uno) de grado mínimo tal que T es un cero de $\mathfrak{m}$; es decir, $\mathfrak{m}(T) = \theta$.

(ii) Demostrar que si p es cualquier otro polinomio y T es un cero de p, entonces $\mathfrak{m}$ divide a p. En particular $\mathfrak{m}$ divide al polinomio característico de T.

(iii) Mostrar que sólo puede existir un polinomio mónico de grado mínimo que se anula en T.

Al polinomio mónico $\mathfrak{m}$ de grado mínimo que se anula en T se le llama el *polinomio mínimo* del operador T.

(b) Sea A una matriz cuadrada.

(i) Probar que existe un polinomio mónico $\mathfrak{m}$ (el coeficiente de la mayor potencia es uno) de grado mínimo tal que T es un cero de $\mathfrak{m}$; es decir, $\mathfrak{m}(A) = \mathscr{O}$.

(ii) Demostrar que si p es cualquier otro polinomio con A es un cero de p; entonces $\mathfrak{m}$ divide a p. En particular $\mathfrak{m}$ divide al polinomio característico de A.

(iii) Mostrar que sólo puede existir un polinomio mónico de grado mínimo que se anula en A.

Al polinomio mónico $\mathfrak{m}$, de grado mínimo y que se anula en A, se le llama el *polinomio mínimo* de la matriz A.

DEMOSTRACIÓN ■ *(a)* *(i)* Por el teorema de Cayley-Hamilton (ejercicio resuelto 50) existe al menos un polinomio que se anula en T, el polinomio característico de T. De entre todos los polinomios que se anulan en T hay al menos uno, $\mathfrak{m}$, de grado mínimo. Es claro que de ser necesario se puede dividir entre el coeficiente principal (el coeficiente de la mayor potencia) y hacer que $\mathfrak{m}(x)$ sea un polinomio mónico.

(ii) Sea $p(x)$ cualquier polinomio tal que $p(T) = \theta$. Por algoritmo de división existen un par de polinomios q, r tales que

$$p(x) = \mathfrak{m}(x)q(x) + r(x)$$

con $r(x) = 0$, el polinomio constante cero, o $0 \leq \text{grad}(r) \leq \text{grad}(\mathfrak{m}) - 1$ ($\text{grad}(w)$ denota el grado del polinomio w). Entonces

$$\begin{aligned}\theta &= p(T)\\ &= \mathfrak{m}(T)q(T) + r(T)\\ &= \theta + r(T)\\ &= r(T)\end{aligned}$$

y por tanto $r(T) = \theta$. Así T se anula en el polinomio $r(x)$ el cual es el polinomio constante cero o tiene grado menor que el grado del polinomio $\mathfrak{m}$; por la definición de $\mathfrak{m}$ la última alternativa no puede ser, entonces se concluye que $r(x)$ es el polinomio constante cero. Luego $\mathfrak{m}$ divide a p.

(iii) Sea $p_1(x)$ un polinomio mónico de grado mínimo que se anula en T. Entonces $\mathfrak{m}$ y p_1 tienen el mismo grado y por el inciso anterior existe un escalar c tal que

$$p_1(x) = c\mathfrak{m}(x)$$

y puesto que ambos son mónicos, se desprende que $c = 1$ y por tanto $p_1(x) = \mathfrak{m}(x)$.

(b) Es consecuencia inmediata del inciso anterior al poner $T = T_A$. ■

53 Demostrar que el polinomio mínimo y el polinomio característico de un operador T, en un espacio de dimensión finita, o de una matriz cuadrada, tienen las mismas raíces y por ende, los mismos factores irreducibles.

DEMOSTRACIÓN ■ Sean $p_T(x)$ y $\mathfrak{m}(x)$ los polinomios característico y mínimo del operador T. Por el ejercicio resuelto 51 de este segmento, toda raíz del polinomio $p_T(x)$ es raíz del polinomio mínimo, pues toda raíz del polinomio característico es valor propio del operador T. Inversamente, puesto que $\mathfrak{m}(x)$ divide al polinomio $p_T(x)$, toda raíz del polinomio mínimo es raíz del polinomio característico. Si λ_i, $i = 1, \ldots, k$, son los valores propios distintos entre sí del operador T, entonces los factores irreducibles p_T son $(x - \lambda_1), \ldots, (x - \lambda_k)$

y, ya que ambos polinomios tienen las mismas raíces, éstos son también los factores irreducibles de $\mathfrak{m}$. El caso para una matriz cuadrada A es consecuencia inmediata de lo precedente al tomar $T = T_A$. ■

54 Encontrar el polinomio mínimo del operador lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido por $T(x,y,z) = (-y, 2x + 3y, -x - y + z)$.

Solución

$$\begin{aligned} T(1,0,0) &= (0,2,-1), \\ T(0,1,0) &= (-1,3,-1), \\ T(0,0,1) &= (0,0,1). \end{aligned}$$

Por tanto,

$$[T]_{\mathscr{B}} = \begin{bmatrix} 0 & -1 & 0 \\ 2 & 3 & 0 \\ -1 & -1 & 1 \end{bmatrix}$$

y

$$\begin{aligned} p_T(\lambda) &= \begin{vmatrix} 0-\lambda & -1 & 0 \\ 2 & 3-\lambda & 0 \\ -1 & -1 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)\begin{vmatrix} -\lambda & -1 \\ 2 & 3-\lambda \end{vmatrix} \\ &= (1-\lambda)(\lambda^2 - 3\lambda + 2) \\ &= -(\lambda-1)(\lambda-1)(\lambda-2) \\ &= -(\lambda-1)^2(\lambda-2). \end{aligned}$$

Puesto que el polinomio mínimo $\mathfrak{m}(\lambda)$ divide a $p_T(\lambda)$, ambos tienen los mismos factores lineales y $\mathfrak{m}$ es mónico, entonces $\mathfrak{m}(\lambda)$ es uno de estos polinomios

$$\begin{aligned} p_1(\lambda) &= (\lambda-1)^2(\lambda-2), \\ p_2(\lambda) &= (\lambda-1)(\lambda-2) = \lambda^2 - 3\lambda + 2. \end{aligned}$$

Y como

$$p_1([T]_{\mathscr{B}}) = -p_T([T]_{\mathscr{B}}) = -\mathscr{O} = \mathscr{O}$$

y

$$\begin{aligned} p_2([T]_{\mathscr{B}}) &= (A-I)(A-2I) \\ &= \begin{bmatrix} -1 & -1 & 0 \\ 2 & 2 & 0 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} -2 & -1 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

se tiene que

$$\mathfrak{m}(\lambda) = p_2(\lambda) = \lambda^2 - 3\lambda + 2$$

es el polinomio mínimo de T. ✔

5.4.2 Ejercicios propuestos

El lector encontrará la respuesta a los ejercicios en cursiva en el apéndice E al final del libro.

Transformaciones lineales (respuestas en páginas 1084-1086)

En los ejercicios 1 a 20 utilizar la definición 5.1 para determinar si la transformación $T: \mathbf{E} \to \mathbf{F}$, definida por la fórmula dada para todo vector en $\mathbf{E}$, es o no lineal. En caso afirmativo probar rigurosamente que se cumplen las dos condiciones de esta definición y en caso contrario mostrar las propiedades que no se cumplen mediante contraejemplos.

1 $T: \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (-x, x+y, 2x-2y)$.

2 $T: \mathbb{R} \to \mathbb{R}^2$, $T(x,y) = (e^x, e^y)$.

3 $T: \mathbb{R} \to \mathbb{R}^2$, $T(x) = (x, e^x)$.

4 $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x,0)$.

5 $T: \mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (x,y,0)$.

6 $T: \mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (0,0,z)$.

7 $T: \mathbb{R}^3 \to \mathbb{R}^2$, $T(x,y,z) = (-x+2y+z, 2x-y+z)$.

8 $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (y,x)$.

9 $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x,-y)$.

10 $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (-x,y)$.

11 $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x^2, y^2)$.

12 $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x,x)$.

13 $T: \mathbb{R}^2 \to \mathbb{R}$, $T(x,y) = xy$.

14 $T: \mathbb{R}^2 \to \mathbb{R}$, $T(x,y) = -3x+2y$.

15 $T: C[0,1] \to \mathbb{R}$, $T(f) = f(0)$.

16 $T: C[0,1] \to \mathbb{R}^2$, $T(f) = (f(0), f(1))$.

17 $T: C^1[-1,1] \to \mathbb{R}^2$, $T(f) = (f'(1/2), \int_{-1}^{1} f(x)dx)$.

18 $T: \mathfrak{M}_n \to \mathbb{R}$, $T(A) = \det(A)$.

19 $T: \mathbf{P} \to \mathbf{P}$, $T(p) = q$, donde $q(x) = p(x-1)$.

20 $T:\mathscr{L}_1[a,\infty)\to\mathbb{R}$, $T(f)=\int_a^\infty f(x)dx$, donde $\mathscr{L}_1[a,\infty)$ es el espacio vectorial del ejercicio resuelto 37 de la página 372.

En los ejercicios 21 a 30 determinar si la transformación $T:\mathbb{R}^n\to\mathbb{R}^m$ es lineal encontrando, si es posible, una representación matricial relativa a las bases canónicas de los espacios $\mathbb{R}^n$ y $\mathbb{R}^m$ como se hizo en el ejemplo 5.11.

21 $T:\mathbb{R}^3\to\mathbb{R}^4$,

$$T(x_1,x_2,x_3)=(-5x_1-x_3,-3x_1-7x_2,2x_2+x_3,-3x_1+3x_2-x_3).$$

22 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(x,y,z)=(x+z,-x+y+z,-2x+y+3z)$

23 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(x,y,z)=(yx+x,xz-y,x+y+z)$.

24 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(x,y,z)=(x,-y,x+y+z)$.

25 $T:\mathbb{R}^2\to\mathbb{R}^3$, $T(x,y)=(2x-3y,-x+y)$.

26 $T:\mathbb{R}^4\to\mathbb{R}^3$,

$$T(x_1,x_2,x_3,x_4)=(x_1-x_2+2x_3-x_4,-2x_1-x_2+x_4,x_1-x_2+x_3-3x_4).$$

27 $T:\mathbb{R}^2\to\mathbb{R}^2$, $T(x,y)=(x-y,x^2-y^2)$.

28 $T:\mathbb{R}^3\to\mathbb{R}$, $T(x,y,z)=-2x+3y-z$.

29 $T:\mathbb{R}^3\to\mathbb{R}^2$, $T(x,y,z)=(x+3y-z,2x-y+2z)$.

30 $T:\mathbb{R}^4\to\mathbb{R}^2$, $T(x_1,x_2,x_3,x_4)=(x_1+x_2-x_3-x_4,x_1+x_2)$.

En los ejercicios 31 a 37 utilizar el ejercicio resuelto 1 de esta sección para encontrar una transformación lineal $T:\mathbb{R}^n\to\mathbb{R}^m$ tal que $T(\vec{e}_i)=\vec{f}_i$.

31 $T:\mathbb{R}^3\to\mathbb{R}^4$, $T(1,0,1)=(-1,1,2,1)$, $T(-1,1,1)=(-2,3,2,4)$ y $T(0,1,1)=(-2,1,3,0)$.

32 $T:\mathbb{R}^2\to\mathbb{R}^3$, $T(1,2)=(1,-1,1)$ y $T(-1,1)=(-2,3,1)$.

33 $T:\mathbb{R}^4\to\mathbb{R}^3$, $T(1,0,1,1)=(1,1,-1)$, $T(0,-1,1,1)=(2,0,1)$, $T(-2,1,2,1)=(1,0,1)$ y $T(-1,0,1,1)=(-2,1,2)$.

34 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(1,2,1)=(1,2,-1)$ y $T(-1,1,0)=(2,-2,1)$ y $T(-1,1,3)=(1,2,-1)$.

35 $T:\mathbb{R}^2\to\mathbb{R}^2$, $T(1,1)=(1,2)$ y $T(-1,1)=(1,2)$.

36 $T:\mathbf{P}_2\to\mathbf{P}_3$, $T(1)=x$, $T(2x)=x^2$, $T(3x^2)=x^3$.

37 $T:\mathfrak{M}_2\to\mathbb{R}^4$, $T\left(\begin{bmatrix}1&0\\0&1\end{bmatrix}\right)=(-1,2,1,1)$, $T\left(\begin{bmatrix}0&1\\1&0\end{bmatrix}\right)=(1,0,1,-1)$, $T\left(\begin{bmatrix}-1&0\\1&1\end{bmatrix}\right)=(2,-1,1,0)$ y $T\left(\begin{bmatrix}1&0\\0&0\end{bmatrix}\right)=(0,0,1,1)$.

38 Si $\mathbf{E}$, $\mathbf{F}$ son dos espacios vectoriales y $T:\mathbf{E}\to\mathbf{E}$ es una transformación, probar que T es lineal si y sólo si

$$T(\alpha\vec{u}+\beta\vec{v})=\alpha T(\vec{u})+\beta T(\vec{v})$$

para todo par de vectores $\vec{u},\vec{v}\in\mathbf{E}$ y para todo par de escalares α,β.

39 Encontrar la aplicación lineal $T:\mathbb{R}^2\to\mathbb{R}^2$ tal que a todo vector $\vec{u}$ lo transforma en el vector que se obtiene al girar ϕ radianes en sentido contrario a las manecillas del reloj a $\vec{u}$ y que tiene la misma norma que este vector.

40 Encontrar la aplicación lineal $T:\mathbb{R}^2\to\mathbb{R}^2$ que a todo vector $\vec{u}$ lo transforma en la reflexión de éste respecto al eje x.

41 Encontrar la aplicación lineal $T:\mathbb{R}^2\to\mathbb{R}^2$ que a todo vector $\vec{u}$ lo transforma en la reflexión de éste respecto al eje y.

42 Encontrar la aplicación lineal $T:\mathbb{R}^3\to\mathbb{R}^3$ tal que a todo vector $\vec{u}$ lo transforma en el vector que se obtiene al girar ϕ radianes en sentido contrario a las manecillas del reloj a $\vec{u}$ alrededor del eje z y que tiene la misma norma que este vector.

En los ejercicios 43 a 50 utilizar los ejercicios resueltos 4 y 5 para encontrar una transformación lineal $T:\mathbf{E}\to\mathbf{F}$ cuya imagen esté generada por los vectores $\vec{f}_i$ del espacio $\mathbf{F}$.

43 $\mathbf{E}=\mathbb{R}^3$, $\mathbf{F}=\mathbb{R}^4$, $\vec{f}_1=(-1,2,1,1)$, $\vec{f}_2=(-1,0,1,1)$ y $\vec{f}_3=(2,-1,0,2)$.

44 $\mathbf{E}=\mathbb{R}^2$, $\mathbf{F}=\mathbb{R}^2$, $\vec{f}_1=(-1,1)$ y $\vec{f}_2=(3,1)$.

45 $\mathbf{E}=\mathbb{R}^3$, $\mathbf{F}=\mathbb{R}^3$, $\vec{f}_1=(1,1,1)$ y $\vec{f}_2=(2-1,1)$.

46 $\mathbf{E}=\mathbb{R}^2$, $\mathbf{F}=\mathbb{R}^4$, $\vec{f}_1=(1,0,1,1)$, $\vec{f}_2=(0,-1,1,1)$ y $\vec{f}_3=(2,1,1,2)$.

47 $\mathbf{E}=$gn$(1,x,x^2)<\mathbf{P}$, $\mathbf{F}=\mathbf{P}$, $\vec{f}_1=x^2-2x$, $\vec{f}_2=x^3+1$.

48 $\mathbf{E}=$gn$(1,x+1,x^2-x)<\mathbf{P}$, $\mathbf{F}=\mathbf{P}$, $\vec{f}_1=1$, $\vec{f}_2=x^2+1$, $\vec{f}_3=3x-2$.

49 $\mathbf{E}=\text{gn}(\cos x,\operatorname{sen}x)<\mathfrak{F}(\mathbb{R})$, $\mathbf{F}=\mathbb{R}^2$, $\vec{f}_1=(1,1)$, $\vec{f}_2=(-1,2)$.

50 $\mathbf{E}$ el espacio de matrices simétricas 2×2, $\mathbf{F}=\mathbb{R}^3$, $\vec{f}_1=(1,0,0)$, $\vec{f}_2=(0,1,0)$ y $\vec{f}_3=(0,0,1)$.

51 Probar que si $\mathbf{E}$ es un espacio de dimensión finita n; $\mathbf{F}$ es un espacio vectorial cualquiera; $\vec{e}_1,\ldots,\vec{e}_m$, con $0<m<n$, son vectores L.I. del espacio $\mathbf{E}$ y $\dim(\mathbf{F})>n-m$; entonces existe $T\in\mathscr{L}(\mathbf{E},\mathbf{F})$, no nula, tal que $\text{Ker}(T)=\text{gn}(\vec{e}_1,\ldots,\vec{e}_m)$.

En los ejercicios 52 a 57 encontrar, utilizando el ejercicio precedente, una transformación $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$ cuyo núcleo esté generado por los vectores dados del espacio $\mathbf{E}$.

52 $\mathbf{E} = \mathbb{R}^4$, $\mathbf{F} = \mathbb{R}^3$, $\vec{e}_1 = (-1,2,1,0)$, $\vec{e}_2 = (1,0,1,1)$.

53 $\mathbf{E} = \mathbb{R}^4$, $\mathbf{F} = \mathbb{R}^3$, $\vec{e}_1 = (-1,1,-1,0)$, $\vec{e}_2 = (-2,0,1,1)$.

54 $\mathbf{E} = \mathbb{R}^3$, $\mathbf{F} = \mathbb{R}^3$, $\vec{e}_1 = (-1,0,1)$.

55 $\mathbf{E} = \mathbb{R}^3$, $\mathbf{F} = \mathbb{R}^4$, $\vec{e}_1 = (1,0,1)$, $\vec{e}_2 = (-1,0,1)$.

56 $\mathbf{E} = \mathbf{P}_3$, $\mathbf{F} = \mathbf{P}$, $\vec{e}_1 = x-1$, $\vec{e}_2 = x$.

57 $\mathbf{E} = \mathfrak{M}_2$, $\mathbf{F} = \mathbb{R}^4$, $\vec{e}_1 = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$, $\vec{e}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

58 Encontrar una transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$ cuyo núcleo esté generado por $(-1,2,1)$ y cuya imagen esté generada por $(-2,1,3)$ y $(-1,0,1)$.

59 Encontrar una transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^4$ cuyo núcleo esté generado por $(-1,0,1)$, $(0,-1,1)$ y cuya imagen esté generada por $(1,-1,0,2)$ y $(3,-1,0,1)$.

60 Hallar una transformación lineal $T : \mathbf{P}_3 \to \mathbf{P}$ cuyo núcleo esté generado por $x-2$, x^2 y cuya imagen esté generada por $x-1$, x^3.

61 Encontrar una transformación lineal $T : \mathbf{P}_4 \to \mathbf{P}$ tal que su núcleo esté generado por $x-1$, $2-x$ y la imagen esté generada por x^2-2, $4-x^2$, x^3+1.

62 Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación (no necesariamente lineal); entonces $T(\vec{u}) = (T_1(\vec{u}), \ldots, T_m(\vec{u}))$, donde $T_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$. A las tranformaciones T_i se les llaman funciones componentes de la aplicación T. Por ejemplo, si $T : \mathbb{R}^3 \to \mathbb{R}^2$ está definida por $T(x,y,z) = (x^2 - yz, xz + y^2)$ para todo $(x,y,z) \in \mathbb{R}^3$, entonces las funciones componentes de T son $T_1(x,y,z) = x^2 - yz$ y $T_2(x,y,z) = xz + y^2$.

(a) Sea $T : \mathbb{R}^n \to \mathbb{R}$ una transformación. Probar que T es lineal si y solo si existen $a_i \in \mathbb{R}$, $i = 1, \ldots, n$, tales que

$$T(x_1, \ldots, x_n) = a_1x_1 + \cdots + a_nx_n.$$

(b) Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación con funciones componentes $T_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$. Probar que T es lineal si y sólo si todas sus funciones componentes son lineales.

(c) Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación con funciones componentes $T_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$. Probar que T es lineal si y sólo si todas sus funciones componentes son de la forma

$$T_j(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_{ji}x_i$$

para ciertos escalares a_{ji}.

En los ejercicios 63 a 77 utilice el ejercicio 62*(c)* para determinar por simple inspección si la transformación $T:\mathbb{R}^n \to \mathbb{R}^m$, definida por la fórmula indicada para todo $\vec{u} \in \mathbb{R}^n$, es lineal o no.

63 $T:\mathbb{R}^3 \to \mathbb{R}^2$, $T(x,y,z) = (x-y, x+y+z)$.

64 $T:\mathbb{R}^3 \to \mathbb{R}^2$, $T(x,y,z) = (x+y, x+y^2+z)$.

65 $T:\mathbb{R}^4 \to \mathbb{R}^3$, $T(x_1,x_2,x_3,x_4) = (0, x_1-x_2, 2x_1+x_2+3x_3-x_4)$.

66 $T:\mathbb{R}^4 \to \mathbb{R}^2$, $T(x_1,x_2,x_3,x_4) = (x_1x_2+x_3, x_1+x_4)$.

67 $T:\mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x\cos\theta, y\operatorname{sen}\theta)$, donde $\theta \in \mathbb{R}$ es dado.

68 $T:\mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (\theta\cos x, \theta\operatorname{sen} y)$, donde $\theta \in \mathbb{R}$ es dado.

69 $T:\mathbb{R}^4 \to \mathbb{R}^4$, $T(x_1,x_2,x_3,x_4) = (x_1+1, x_2+1, x_3+1, x_4+1)$.

70 $T:\mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (x, 2y, 3z)$.

71 $T:\mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x, y^2)$.

72 $T:\mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x, e^y)$.

73 $T:\mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x, |y|)$.

74 $T:\mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (y, -x, z)$.

75 $T:\mathbb{R}^n \to \mathbb{R}$, $T(x_1,\ldots,x_n) = \frac{1}{n}\sum_{i=1}^n x_i$.

76 $T:\mathbb{R}^n \to \mathbb{R}$, $T(x_1,\ldots,x_n) = \frac{1}{n}\prod_{i=1}^n x_i = \frac{1}{n}x_1,\ldots,x_n$.

77 $T:\mathbb{R}^n \to \mathbb{R}^n$, $T(x_1,\ldots,x_n) = (x_1, x_1+x_2, \ldots, x_1+\cdots+x_n)$.

78 Sean $\mathbf{E}$ un espacio con producto interior $\langle\cdot,\cdot\rangle$, $\vec{v} \in \mathbf{E} - \{\vec{0}_\mathbf{E}\}$ un vector dado, $S = \operatorname{gn}(\vec{v})$ y $\pi:\mathbf{E} \to S$ definida por $\pi(\vec{u}) = p_{\vec{u}}$ donde $\vec{p}_{\vec{u}}$ es el vector proyección de cada $\vec{u} \in \mathbf{E}$ sobre $\vec{v}$.

(a) Probar que π es lineal.

(b) Determinar el núcleo de π.

79 Sean $\mathbf{E} = \mathbb{R}^n$, $\vec{e}_1,\ldots,\vec{e}_m \in \mathbf{E}$ vectores L.I. dados, $S = \operatorname{gn}(\vec{e}_1,\ldots,\vec{e}_m)$ y $T:\mathbf{E} \to S$ definida por $\pi(\vec{u}) = p_{\vec{u}}$ donde $\vec{p}_{\vec{u}}$ es el vector proyección de cada $\vec{u} \in \mathbf{E}$ sobre S.

(a) Probar que π es lineal.

(b) Determinar el núcleo de π.

80 Sean $\mathbf{E}$ un espacio con producto interior $\langle\cdot,\cdot\rangle$, $S \neq \{\vec{0}_\mathbf{E}\}$ un subespacio de dimensión finita de $\mathbf{E}$ y $\pi:\mathbf{E} \to S$ definida por $\pi(\vec{u}) = p_{\vec{u}}$ donde $\vec{p}_{\vec{u}}$ es el vector proyección de cada $\vec{u} \in \mathbf{E}$ sobre S.

(a) Probar que π es lineal.

(b) Determinar el núcleo de π.

En los ejercicios 81 a 90 $T:\mathbb{R}^n \to \mathbb{R}^m$ es una transformación definida por la fórmula dada para todo $\vec{u} \in \mathbb{R}^n$. *(i)* Por simple inspección, utilizando el ejercicio 62*(c)*, verificar que T es lineal; *(ii)* hallar una matriz $A \in \mathfrak{M}_{m\times n}$ tal que $T = T_A$ (cfr. la discusión dada en la pág. 420); *(iii)* hallar una base y $\dim(T(\mathbb{R}^n))$; *(iv)* encontrar una base y la dimensión de $\operatorname{Ker}(T)$; *(v)* determinar si T es inyectiva y *(vi)* determinar si T es suprayectiva.

81 $T:\mathbb{R}^3 \to \mathbb{R}^2$, $T(x,y,z) = (-x+2y-3z, x+y-z)$.

82 $T:\mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x+y,-y)$.

83 $T:\mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (-x+y+z, -2x+y+3z, x-2z)$.

84 $T:\mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (x+y,y,0)$.

85 $T:\mathbb{R}^4 \to \mathbb{R}^3$, $T(x_1,x_2,x_3,x_4) = (x_1-x_2+3x_3, -x_2+x_1, -x_1+x_2-4x_3+x_4)$.

86 $T:\mathbb{R}^4 \to \mathbb{R}^4$, $T(x_1,x_2,x_3,x_4) = (x_1, x_2+x_3, x_1-x_4, x_1+x_2+x_3)$.

87 $T:\mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (x-y+z, 2x-y-3z, 3x-y-7z)$.

88 $T:\mathbb{R}^3 \to \mathbb{R}^4$, $T(x,y,z) = (x-y+z, x+3y-4z, x-2y+z, 3x-4y-z)$.

89 $T:\mathbb{R}^4 \to \mathbb{R}^4$, $T(x_1,x_2,x_3,x_4) = (x_1-x_2, x_1+x_3-x_4, -x_1-2x_2-3x_3+3x_4, -x_2-x_3)$.

90 $T:\mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (x-y+z, -2x+y+z, x-y+2z)$.

91 Sea $T \in \mathscr{L}(\mathbb{R}^n,\mathbb{R})$, probar que T es suprayectiva o T es la aplicación constante cero.

92 Sea $T \in \mathscr{L}(\mathbb{R},\mathbb{R}^n)$, probar que T es inyectiva o T es la aplicación constante cero.

93 Mostrar que si T_1 y T_2 son transformaciones lineales de un espacio vectorial $\mathbf{E}$ de dimensión finita en un espacio $\mathbf{F}$ y $\{\vec{u}_1,\ldots,\vec{u}_n\}$ es una base de $\mathbf{E}$, entonces $T_1 = T_2$ si y sólo si $T_1(\vec{u}_i) = T_2(\vec{u}_i)$ para todo $i = 1,\ldots,n$.

94 Sean $\mathbf{E}$, $\mathbf{F}$ dos espacios vectoriales con $\dim(\mathbf{E})$ finita, $S < \mathbf{E}$ y $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$. Demostrar que

$$\dim(S) = \dim(S \cap \operatorname{Ker}(T)) + \dim(T(S)).$$

95 Sean $\mathbf{E}$, $\mathbf{F}$ espacios vectoriales y $T \in \mathscr{L}(\mathbf{E},\mathbf{F})$. Mostrar que si $T(\vec{u}_1),\ldots,T(\vec{u}_m)$ son L.I. en $\mathbf{F}$, entonces $\vec{u}_1,\ldots,\vec{u}_m$ son L.I. en $\mathbf{E}$. ¿Es cierto el recíproco? Si la respuesta es negativa, dar un contraejemplo.

96 Sean $\mathbf{E}$ y $\mathbf{F}$ un par de espacios vectoriales y $T_1, T_2 : \mathbf{E} \to \mathbf{F}$ dos transformaciones lineales tales que $T_1(\mathbf{E}) \cap T_2(\mathbf{E}) = \{\vec{0}_{\mathbf{F}}\}$, mostrar que T_1 y T_2 son L.I. en el espacio vectorial $\mathscr{L}(\mathbf{E},\mathbf{F})$ (cfr. el ejercicio resuelto 31 del capítulo 3, pág. 190).

97 Sean $\mathbf{P}$ el espacio de polinomios y $n \geq 1$ un entero. Para cada $k = 1,\ldots,n$ sea T_k el operador lineal en $\mathbf{P}$ definido por $T_k(p) = p^{(k)}$, donde $p^{(k)}$ es la k-ésima derivada del polinomio p. Demostrar que $T_1,\ldots,T_n$ son L.I.

98 Probar que si $T \in \mathscr{L}(\mathbf{E}, \mathbf{F})$, entonces

$$\dim(T(\mathbf{E})) \leq \dim(\mathbf{F}).$$

99 Sean $A \in \mathfrak{M}_{m\times n}$ y $B \in \mathfrak{M}_{n\times p}$, probar que

$$\operatorname{Rang}(AB) \leq \operatorname{Rang}(B) \quad \text{y}$$
$$\operatorname{Rang}(AB) \leq \operatorname{Rang}(A).$$

100 Sea $T : \mathbb{R}^\infty \to \mathbb{R}^\infty$ la función definida por $T((a_n)) = (a_{2n})$ para cada sucesión $(a_n) \in \mathbb{R}^\infty$.

(a) Demostrar que T es un operador lineal en el espacio de sucesiones $\mathbb{R}^\infty$.

(b) Probar que T es suprayectivo.

(c) Demostrar que T es singular (no inyectivo).

(d) Sea c_{00} el subespacio de sucesiones finitas (cfr. el ejercicio propuesto 357 del cap. 3, pág. 228). Demostrar que $T(c_{00}) = c_{00}$.

(e) Mostrar que (cfr. el ejercicio resuelto 8 de esta sección)

$$c_{00} \subset T^{-1}(c_{00})$$

pero $c_{00} \neq T^{-1}(c_{00})$.

101 Sea $T : C[0,1] \to C[0,1]$ el operador definido por $T(f) = g$ donde $g(x) = \int_0^x f(t)dt$, mostrar que T es inyectivo pero no suprayectivo.

102 Sean $\mathbf{E}$ y $\mathbf{F}$ un par de espacios vectoriales con dimensión de $\mathbf{E}$ finita, $V < \mathbf{F}$, $T \in \mathscr{L}(\mathbf{E}, \mathbf{F})$ y $T^{-1}(V)$ la imagen inversa del subespacio V (cfr. el ejercicio resuelto 8). Probar que

$$\dim(T^{-1}(V)) = \dim(\operatorname{Ker}(T) + \dim(V \cap T(\mathbf{E})).$$

En los ejercicios 103 a 107 $\mathbf{E} = \mathfrak{M}_n$ es el espacio de matrices de orden n y $T : \mathbf{E} \to \mathbf{E}$ es el operador definido como

$$T(A) = \frac{A + A^t}{2}$$

para toda $A \in \mathbf{E}$.

103 Probar que T es un operador lineal en $\mathbf{E}$.

104 Probar que $\operatorname{Ker}(T)$ es el subespacio S_2 matrices antisimétricas (cfr. el ejercicio resuelto 28, pág. 187).

105 Demostrar que $T(\mathbf{E})$ es el subespacio de todas las matrices simétricas (cfr. el ejemplo 3.25).

106 En el ejercicio propuesto 353 del capítulo 3 (pág. 227), se pide encontrar la dimensión del subespacio de matrices simétricas. Utilizar ese resultado para calcular la dimensión del subespacio de matrices antisimétricas.

107 Sea $L : \mathbf{E} \to \mathbf{E}$ el operador definido, para cada $A \in \mathbf{E}$, por

$$L(A) = \frac{A - A^t}{2}.$$

(a) Probar que L es lineal.

(b) Describir $\text{Ker}(L)$, encontrar una base para este subespacio y su dimensión.

(c) Determinar $L(\mathbf{E})$, encontrar un base y la dimensión de este subespacio.

108 Sean $\mathscr{F}(\mathbb{R})$ el espacio de funciones reales con dominio en $\mathbb{R}$ y $T : \mathscr{F}(\mathbb{R}) \to \mathscr{F}(\mathbb{R})$ el operador definido, para cada $f \in \mathscr{F}(\mathbb{R})$, por $T(f) = g$ donde

$$g(x) = \frac{f(x) + f(-x)}{2}$$

para todo $x \in \mathbb{R}$.

(a) Probar que T es lineal.

(b) Determinar el núcleo y la imagen de T (cfr. los ejercicios 124 y 125 del capítulo 3).

109 Sean $\mathscr{F}(\mathbb{R})$ el espacio de funciones reales con dominio en $\mathbb{R}$ y $T : \mathscr{F}(\mathbb{R}) \to \mathscr{F}(\mathbb{R})$ el operador definido, para cada $f \in \mathscr{F}(\mathbb{R})$, por $T(f) = g$ donde

$$g(x) = \frac{f(x) - f(-x)}{2}$$

para todo $x \in \mathbb{R}$.

(a) Probar que T es lineal.

(b) Determinar el núcleo y la imagen de T (cfr. los ejercicios 124 y 125 del capítulo 3).

En los ejercicios 110 a 124 se define una transformación T entre los espacios vectoriales indicados $\mathbf{E}$ y $\mathbf{F}$. *(i)* Probar que T es lineal; *(ii)* determinar el núcleo y la imagen de T; *(iii)* calcular las dimensiones de $\text{Ker}(T)$ y $T(\mathbf{E})$ y en caso de ser finitas hallar bases para estos subespacios.

110 $\mathbf{E} = \mathbf{F} = \mathbf{P}_n$, $T : \mathbf{E} \to \mathbf{F}$ definida por $T(p) = q$ donde, para cada $p \in \mathbf{E}$, $q(x) = p(x-1)$ para todo x.

111 $\mathbf{E} = \mathbf{F} = \mathbf{P}_3$, $T : \mathbf{E} \to \mathbf{F}$ definida por $T(f) = g$ donde, para cada $f \in \mathbf{E}$, $g(x) = xf'(x)$ para todo x.

112 $\mathbf{E} = \mathbf{F} = \mathbf{P}_3$, $T : \mathbf{E} \to \mathbf{F}$ definida por $T(f) = g$ donde, para cada $f \in \mathbf{E}$, $g(x) = f'(x)$ para todo x.

113 $\mathbf{E} = \mathbf{P}$, $\mathbf{F} = \mathbb{R}$, $T : \mathbf{E} \to \mathbf{F}$ definida por $T(f) = f(0)$.

114 $\mathbf{E} = \mathbf{P}_3$, $\mathbf{F} = \mathbb{R}$, $T : \mathbf{E} \to \mathbf{F}$ definida por $T(f) = f(0)$.

115 $\mathbf{E} = \mathfrak{M}_2$, $\mathbf{F} = \mathbb{R}$, $T : \mathbf{E} \to \mathbf{F}$, $T(A) = \text{tra}(A)$, la traza de la matriz A (la suma de los elementos de la diagonal).

116 $\mathbf{E} = \mathfrak{M}_3$, $\mathbf{F} = \mathbb{R}$, $T : \mathbf{E} \to \mathbf{F}$, $T(A) = \text{tra}(A)$, la traza de la matriz A (la suma de los elementos de la diagonal).

117 $\mathbf{E} = \mathbf{P}_2$, $\mathbf{F} = \mathbf{P}_3$, $T : \mathbf{E} \to \mathbf{F}$, $T(p) = q$, donde $q(x) = xp(x) + p'(x)$.

118 $\mathbf{E} = C^1(\mathbb{R})$, el espacio de funciones con valores reales derivables con continuidad en $\mathbb{R}$; $\mathbf{F} = C(\mathbb{R})$, el espacio de funciones con valores reales continuas en $\mathbb{R}$; $T : \mathbf{E} \to \mathbf{F}$ definida por $T(f) = g(x)$, donde $g(x) = xf(x) + f'(x)$ para todo $x \in \mathbb{R}$.

119 $\mathbf{E} = \mathbf{F} = \mathbf{P}_3$, $T : \mathbf{E} \to \mathbf{F}$ definida por $T(f) = g$ donde, para cada $f \in \mathbf{E}$, $g(x) = xf'(x)$ para todo x.

120 $\mathbf{E} = \mathbf{F} = C[-1,1]$, $T : \mathbf{E} \to \mathbf{F}$ definida por $T(f) = g$ donde, para cada $f \in \mathbf{E}$, $g(x) = xf'(x)$ para todo $x \in [-1,1]$.

121 $\mathbf{E} = \mathbf{F} = C[a,b]$, $T : \mathbf{E} \to \mathbf{F}$ definida por $T(f) = g$ donde, para cada $f \in \mathbf{E}$,

$$g(x) = \int_a^b f(t)e^{x-t}dt$$

para todo $x \in [a,b]$. En este caso se debe primero probar que efectivamente $g \in C[a,b]$.

122 $\mathbf{E} = \mathbf{F} = C[a,b]$, $T : \mathbf{E} \to \mathbf{F}$ definida por $T(f) = g$ donde, para cada $f \in \mathbf{E}$,

$$g(x) = \int_a^b f(t)\,\mathrm{sen}(x-t)dt$$

para todo $x \in [a,b]$. En este caso se debe primero probar que efectivamente $g \in C[a,b]$.

123 $\mathbf{E} = C^2[a,b]$, $\mathbf{F} = C[a,b]$, $T : \mathbf{E} \to \mathbf{F}$ definida por $T(f) = f'' + 2f' - 3f$ para cada $f \in C^2[a,b]$.

124 $\mathbf{E} = \mathbf{F}$ el espacio de sucesiones convergentes (cfr. el ejercicio 112 de la pág. 213), $T : \mathbf{E} \to \mathbf{F}$ definida, para cada $(a_n) \in \mathbf{E}$, por $T((a_n)) = (y_n)$ donde $y_n = a - a_n$ para todo n y $a = \mathrm{lím}_{n\to\infty} a_n$.

En los ejercicios 125 a 129 $\mathbf{E} = C[-\pi,\pi]$, S es el subconjunto de $\mathbf{E}$ de funciones f tales que

$$\int_{-\pi}^{\pi} f(x)dx = 0,\ \int_{-\pi}^{\pi} f(x)\cos x dx = 0 \text{ y } \int_{-\pi}^{\pi} f(x)\,\mathrm{sen}\, x dx = 0.$$

125 Probar que S es un subespacio de $\mathbf{E}$.

126 Probar que las funciones $f_n(x) = \cos(nx)$ y $g_n(x) = \mathrm{sen}(nx)$ pertenecen a S para todo $n = 2,3,\ldots$

127 Probar que S tiene dimensión infinita.

128 Sea $T : \mathbf{E} \to \mathbf{E}$ el operador definido, para cada $f \in \mathbf{E}$, por $T(f) = g(x)$ donde

$$g(x) = \int_{-\pi}^{\pi} (1 + \cos(x-t))f(t)dt$$

para todo $x \in [-\pi,\pi]$, probar que T es lineal.

129 Sea T el operador lineal definido en el ejercicio anterior.

(a) Probar que $T(\mathbf{E})$ tiene dimensión finita, hallar una base para este subespacio y el rango de T (la dimensión de $T(\mathbf{E})$).

(*b*) Encontrar el núcleo de T.

(*c*) Hallar todas las funciones $f \in C[-\pi,\pi]$ tales que $T(f) = \lambda f$ para alguna constante $\lambda \in \mathbb{R}$ (las funciones propias de T).

130 Sean $\mathbf{E}$ un espacio vectorial, T un operador lineal en él que es inyectivo y $\mathbf{E} = T(\mathbf{E})$. Probar que T es una transformación lineal biyectiva del espacio $\mathbf{E}$ al espacio $\mathbf{E}$ y que por tanto es invertible.

En los ejercicios 131 a 137 se define un operador $T : \mathbb{R}^n \to \mathbb{R}^n$en el espacio $\mathbb{R}^n$ indicado. *(i)* Probar que T es un operador lineal; *(ii)* mostrar que T es biyectivo; *(iii)* encontrar $T^{-1}(\vec{u})$ para todo $\vec{u} \in \mathbb{R}^n$.

131 $\mathbb{R}^2$, $T(x,y) = (x+2y, x+y)$.

132 $\mathbb{R}^2$, $T(x,y) = (x-2y, -x+3y)$.

133 $\mathbb{R}^2$, $T(x,y) = (x-y, 2x-3y)$

134 $\mathbb{R}^3$, $T(x,y,z) = (x-y+z, x-2y-z, -x+3y-2z)$.

135 $\mathbb{R}^3$, $T(x,y,z) = (x+y, x-y-z, -x+y-z)$.

136 $\mathbb{R}^3$, $T(x,y,z) = (x+y+3z, x-y, x+z,)$.

137 $\mathbb{R}^3$, $T(x,y,z) = (x+y, 3x+y+z, 2x-y+z)$.

En los ejercicios 138 a 148, $\mathbf{E}$ es un espacio vectorial y $T, T_1, T_2 : \mathbf{E} \to \mathbf{E}$ son operadores lineales. Se acostumbra escribir TT_1 en lugar de la notación empleada para la composición $T \circ T_1$ de estos operadores (cfr. el ejercicio resuelto 12 de esta sección). Se definen $T^0 = I$, donde $I(\vec{u}) = \vec{u}$ para todo $\vec{u} \in \mathbf{E}$ es el operador lineal identidad; $T^n = \underbrace{T \circ \cdots \circ T}_{n}$ o, de manera más rigurosa: $T^1 = T$ y por inducción $T^n = T \circ T^{n-1}$ para $n = 1, 2, \ldots$ y, si T es además invertible, se define $T^{-n} = (T^{-1})^n$ para todo $n = 0, 1, 2, \ldots$ Del ejercicio propuesto 12 de esta sección TT_1, T^n son operadores lineales también.

138 Probar que $T(\alpha T_1) = \alpha(TT_1)$ para todo $\alpha \in \mathbb{R}$.

139 Mostrar que $(\alpha T)(\beta T_1) = \alpha\beta(TT_1)$ para todo par de números reales $\alpha, \beta \in \mathbb{R}$.

140 Demostrar que $T^n T^m = T^{n+m}$ para todo par de números enteros no negativos n, m; y que esta misma igualdad es válida para todo par de números enteros n, m si además el operador T es invertible.

141 Mostrar que $(T^n)^m = T^{nm}$ para todo par de números enteros no negativos n, m y que esta misma igualdad es válida para todo par de números enteros n, m si además el operador T es invertible.

142 Demostrar que

$$T(T_1 + T_2) = TT_1 + TT_2$$

y

$$(T_1 + T_2)T = T_1T + T_2T.$$

143 Probar que $TI = IT = T$.

144 Dar un ejemplo de un espacio $\mathbf{E}$ y un par de operadores lineales T_1 y T_2 tales que $T_1T_2 \neq T_2T_1$.

145 Si T_1 y T_2 conmutan; es decir, $T_1T_2 = T_2T_1$, mostrar que $(T_1T_2)^n = T_1^nT_2^n$ para cualquier entero $n \geq 0$.

146 Si T_1 y T_2 son invertibles, demostrar que T_1T_2 es también invertible y $(T_1T_2)^{-1} = T_2^{-1}T_1^{-1}$; i.e., $(T_1 \circ T_2)^{-1} = T_2^{-1} \circ T_1^{-1}$.

147 Si T_1 y T_2 conmutan y son invertibles, probar que sus respectivos operadores inversos también conmutan.

148 Demostrar que si T_1 y T_2 conmutan entonces:

(a) $(T_1 + T_2)^2 = T_1^2 + 2T_1T_2 + T_2^2$.

(b) $(T_1 + T_2)^3 = T_1^3 + 3T_1^2T_2 + 3T_1T_2^2 + T_2^3$.

(c) Indicar cómo deben modificarse las fórmulas de los dos incisos anteriores en el caso de que $T_1T_2 \neq T_2T_1$.

149 Sean $T_1, T_2 : \mathbb{R}^3 \to \mathbb{R}^3$ los operadores lineales definidos por $T_1(x,y,z) = (x,z,y)$ y $T_2(x,y,z) = (x, x+y, x+y+z)$ para cada $(x,y,z) \in \mathbb{R}^3$.

(a) Encontrar la imagen del punto (x,y,z) al aplicarle cada uno de los operadores lineales *(i)* T_1T_2; *(ii)* T_2T_1; *(iii)* T_1^2; *(iv)* T_2^2; *(v)* $(T_1T_2 - T_2T_1)^2$.

(b) Demostrar que T_1 y T_2 son invertibles y encontrar *(i)* $T_1^{-1}(x,y,z)$, *(ii)* $T_2^{-1}(x,y,z)$; *(iii)* $(T_1T_2)^{-1}(x,y,z)$; *(iv)* $(T_2T_1)^{-1}(x,y,z)$; para todo $(x,y,z) \in \mathbb{R}^3$.

(c) Encontrar $(T - I)^n(x,y,z)$ para todo $(x,y,z) \in \mathbb{R}^3$ y para cada $n \geq 1$.

150 Sea $\mathbf{E}$ un espacio vectorial. Un proyector en $\mathbf{E}$ es un operador lineal $T : \mathbf{E} \to \mathbf{E}$ tal que $T^2 = T$. Probar que si $\mathbf{E}$ es un espacio con producto interior $\langle \cdot, \cdot \rangle$ y S es un subespacio no trivial de dimensión finita de $\mathbf{E}$, entonces $\pi : \mathbf{E} \to \mathbf{E}$, con $\pi(\vec{u})$ el vector proyección de $\vec{u}$ sobre S (cfr. el ejercicio propuesto 80 de esta sección), es un proyector en $\mathbf{E}$.

151 Sean $\mathbf{E}$ un espacio vectorial y S_1 un subespacio de él. Si $T : \mathbf{E} \to \mathbf{E}$ es una proyección sobre S_1 (cfr. el ejercicio 17 de esta sección), mostrar que T es un proyector en $\mathbf{E}$ (cfr. el ejercicio anterior).

152 Si $\mathbf{E}$ es un espacio vectorial y T es un operador lineal en él, mostrar que $T^2 = \theta$, el operador lineal constante cero, si y sólo si $T(\mathbf{E}) \subset \text{Ker}(T)$.

153 Sean $\mathbf{E}$ un espacio vectorial de dimensión finita, $T : \mathbf{E} \to \mathbf{E}$ un operador lineal tal que $\text{Rang}(T^2) = \text{Rang}(T)$.

(a) Mostrar que $T(\mathbf{E}) \cap \text{Ker}(T) = \{\vec{0}_{\mathbf{E}}\}$.

(b) Probar que $\mathbf{E} = T(\mathbf{E}) \oplus \text{Ker}(T)$ (cfr. el ejercicio resuelto 27 del capítulo 3).

154 Sean $\mathbf{P}$ el espacio de polinomios y $D, J : \mathbf{P} \to \mathbf{P}$ los operadores derivación, $D(p) = p'$, y el operador integración, $J(p) = q(x)$ donde $q(x) = \int_0^x p(t)\,dt$ para todo $x \in \mathbb{R}$.

(a) Mostrar que $DJ = I$, el operador identidad en **P**; pero $JD \neq I$.

(b) Hallar Ker(JD) y $(JD)(\mathbf{P})$.

(c) ¿Son invertibles estos operadores? (Cfr. el ejercicio resuelto 14 de esta sección.)

155 Sean **P** el espacio de polinomios, D el operador derivación en **P** y $T : \mathbf{P} \to \mathbf{P}$ el operador definido por $T(p) = q$ donde $q(x) = xp'(x)$ para cada $p \in \mathbf{P}$.

(a) Si $p(x) = 1 - 2x + 3x^2 + x^3$ calcular la imagen de p bajo los operadores *(i)* D, *(ii)* T, *(iii)* DT, *(iv)* TD, *(v)* $DT - TD$, *(vi)* $T^2D^2 - D^2T^2$.

(b) Encontrar los polinomios p para los cuales $T(p) = p$.

(c) Hallar los polinomios p para los cuales $(DT - 2D)(p) = \theta$, el operador constante cero.

(d) Hallar los polinomios p para los cuales $(DT - TD)^n(p) = D^n(p)$, $n \geq 2$ un entero.

156 Sean **P** el espacio de polinomios; $D(p) = p'$ el operador derivación; $T(p) = q$, donde $q(x) = xp(x)$ para cada $p \in \mathbf{P}$; e I el operador identidad en **P**, $I(p) = p$.

(a) Probar que T es un operador lineal en **E**.

(b) Demostrar que $DT - TD = I$.

(c) Mostrar que $DT^n - T^nD = nT^{n-1}$ para cada entero $n \geq 1$.

157 Sean **E** un espacio vectorial; T_1, T_2 un par de operadores lineales en **E**; e I el operador identidad en **E**, $I(\vec{u}) = \vec{u}$ para todo $\vec{u} \in \mathbf{E}$. Si $T_1T_2 - T_2T_1 = I$, probar que

$$T_1T_2^n - T_2^nT_1 = nT_2^{n-1}$$

para todo entero $n \geq 1$.

158 Sean **E** un espacio vectorial, T un operador lineal en él, $T^2 = T \circ T$ el operador composición de T consigo mismo (cfr. el ejercicio resuelto 12 de esta sección) e I el operador lineal identidad ($I(\vec{u}) = \vec{u}$ $\forall \vec{u} \in \mathbf{E}$). Si $T^2 = \theta$, el operador constante cero, probar que $I - T$ es un operador lineal no singular (invertible). (Sugerencia: Utilice el ejercicio resuelto 14 de esta sección y el ejercicio precedente.)

159 Sea T un operador lineal en un espacio vectorial **E** tal que $T^2 + 2T + I = \theta$, el operador constante cero; mostrar que T es invertible.

160 Sean **E** un espacio vectorial, T un operador lineal en él, $T^3 = T \circ T \circ T$ el operador composición de T consigo mismo tres veces (cfr. el ejercicio resuelto 12 de esta sección) e I el operador lineal identidad ($I(\vec{u}) = \vec{u}$ $\forall \vec{u} \in \mathbf{E}$). Si $T^3 = \theta$, el operador constante cero, probar que $I - T$ es un operador lineal no singular (invertible).

161 Probar que si **E** es un espacio vectorial, $T : \mathbf{E} \to \mathbf{E}$ es un operador lineal, $\vec{u} \in \mathbf{E}$ es un vector tal que $T^2(\vec{u}) = \vec{0}_{\mathbf{E}}$ y $T(\vec{u}) \neq \vec{0}_{\mathbf{E}}$, entonces $\vec{u}$ y $T(\vec{u})$ son L.I.

162 Sean **E** un espacio vectorial no nulo, T un operador en él tal que $T^2 = T$, $T \neq \theta$, el operador constante cero, y $T \neq I$, el operador identidad. Mostrar que:

(a) Existe $\vec{u} \in \mathbf{E} - \{\vec{0}_\mathbf{E}\}$ tal que $T(\vec{u}) = \vec{u}$.

(b) Existe $\vec{v} \in \mathbf{E} - \{\vec{0}_\mathbf{E}\}$ tal que $T(\vec{u}) = \vec{0}_\mathbf{E}$.

163 Sean $\mathbf{E}$ y T como en el ejercicio precedente y

$$S_1 = \{\vec{u} \in \mathbf{E} \,|\, T(\vec{u}) = \vec{u}\},$$
$$S_2 = \{\vec{u} \in \mathbf{E} \,|\, T(\vec{u}) = \vec{0}_\mathbf{E}\}.$$

Mostrar que:

(a) $S_1 < \mathbf{E}$ y $S_2 < \mathbf{E}$.

(b) $S_1 \neq \{\vec{0}_\mathbf{E}\}$ y $S_2 \neq \{\vec{0}_\mathbf{E}\}$.

(c) $\mathbf{E} = S_1 \oplus S_2$, la suma directa de S_1 con S_2 (cfr. el ejercicio resuelto 27 del capítulo 3).

164 Sea T un operador lineal en un espacio vectorial $\mathbf{E}$ tal que $T^2 = I$. Suponer que existe $\vec{u} \in \mathbf{E} - \{\vec{0}_\mathbf{E}\}$ tal que $T(\vec{u}) = \alpha\vec{u}$ para algún $\alpha \in \mathbb{R}$. Encontrar los posibles valores de α.

165 Sean $\mathbf{E}$ un espacio vectorial no nulo y T un operador lineal en él tal que $T^2 - 2T + I = \theta$, el operador lineal constante cero. Mostrar que existe $\vec{u} \in \mathbf{E}$, con $\vec{u} \neq \vec{0}_\mathbf{E}$ tal que $T(\vec{u}) = \vec{u}$.

166 Sean $\mathbf{E}$ un espacio vectorial y T un operador lineal en él tal que $T^2 - 2T + I = \theta$, el operador lineal constante cero. Probar que T es invertible y hallar T^{-1}.

167 Sean $\mathbf{E}$ un espacio vectorial, T un operador lineal en él y $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, n$, escalares con $\alpha_n \neq 0$. Probar que si T satisface

$$T^n + \alpha_1 T^{n-1} + \cdots + \alpha_{n-1} T + \alpha_n I = 0$$

entonces T es invertible. Encontrar T^{-1}.

168 Sean $\mathbf{E}$ un espacio vectorial y T un operador lineal en él. T es nilpotente (o nihilpotente) si existe un entero no negativo n tal que $T^n = \theta$, el operador constante cero. Al menor entero no negativo ν tal que $T^\nu = \theta$ se le llama el índice de nilpotencia (o nihilpotencia) de T. Probar que si $\vec{u} \in \mathbf{E}$ es tal que $T^{\nu-1}(\vec{u}) \neq \vec{0}_\mathbf{E}$, entonces los vectores $\vec{u}, T(\vec{u}), T^2(\vec{u}), \ldots, T^{\nu-1}(\vec{u})$ son L.I.

169 Sean $\mathbf{E}$ un espacio vectorial, T un operador nilpotente definido en $\mathbf{E}$ y ν el índice de nilpotencia de él (cfr. el ejercicio precedente).

(a) Probar que $T^m = \theta$ para todo $m \geq \nu$.

(b) Si $\dim(\mathbf{E}) = n < \infty$, demostrar que $T^n = \theta$.

170 Sea el operador lineal $T : \mathbb{R}^4 \to \mathbb{R}^4$ definido por $T(\vec{x}) = A\vec{x}$ donde

$$A = \begin{bmatrix} 5 & 1 & -6 & -3 \\ -2 & 1 & 5 & 3 \\ 9 & 1 & -13 & -7 \\ -12 & -3 & 14 & 7 \end{bmatrix}.$$

Demostrar que T es nilpotente encontrando su índice de nilpotencia (cfr. el ejercicio propuesto 168 de esta sección).

Representación matricial (respuestas en páginas 1086-1090)

En los ejercicios 171 a 186 probar que $\mathscr{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ es una base del espacio $\mathbf{E}$ y hallar $[\vec{u}]_{\mathscr{B}} \in \mathbb{R}^n$, el vector de coordenadas para el vector dado $\vec{u}$ relativo a la base ordenada $(\vec{e}_1, \ldots, \vec{e}_n)$.

171 $\mathscr{B} = \{(1,2),(1,1)\}$, $\mathbf{E} = \mathbb{R}^2$, $\vec{u} = (-1,1)$.

172 $\mathscr{B} = \{(1,3),(1,2)\}$, $\mathbf{E} = \mathbb{R}^2$, $\vec{u} = (2,3)$.

173 $\mathscr{B} = \{(-2,2),(-1,0)\}$, $\mathbf{E} = \mathbb{R}^2$, $\vec{u} = (3,1)$.

174 $\mathscr{B} = \{(3,-2),(-1,4)\}$, $\mathbf{E} = \mathbb{R}^2$, $\vec{u} = (5,-1)$.

175 $\mathscr{B} = \{(1,0,1),(-1,1,0),(1,-1,1)\}$, $\mathbf{E} = \mathbb{R}^3$, $\vec{u} = (-1,3,1)$.

176 $\mathscr{B} = \{(1,2,1),(-1,1,1),(-1,1,2)\}$, $\mathbf{E} = \mathbb{R}^3$, $\vec{u} = (1,0,1)$.

177 $\mathscr{B} = \{(1,-1,3),(-1,2,-3),(2,-2,5)\}$, $\mathbf{E} = \mathbb{R}^3$, $\vec{u} = (3,-3,1)$.

178 $\mathscr{B} = \{(1,-2,-2),(-2,3,1),(-1,2,1)\}$, $\mathbf{E} = \mathbb{R}^3$, $\vec{u} = (2,1,-2)$.

179 $\mathscr{B} = \{(1,-2,3,4),(-2,5,-6,-8),(1,-2,4,4),(3,-5,9,11)\}$, $\mathbf{E} = \mathbb{R}^4$, $\vec{u} = (2,-3,0,1)$.

180 $\mathscr{B} = \{(1,2,-1,0),(-1,-2,9,-1),(2,1,1,1),(-3,-1,1,-2)\}$, $\mathbf{E} = \mathbb{R}^4$, $\vec{u} = (-1,0,1,1)$.

181 $\mathscr{B} = \{x-1, x-2, x^2+1\}$, $\mathbf{E} = \mathbf{P}_2$, $\vec{u} = x^2 - 2x + 3$.

182 $\mathscr{B} = \{1, x-1, x^2-2, x^3+3\}$, $\mathbf{E} = \mathbf{P}_3$, $\vec{u} = 2x^3 - x^2 + 3x$.

183 $\mathscr{B} = \{x-1, x-2, x^2-1\}$, $\mathbf{E} = \mathbf{P}_2$, $\vec{u} = x^2 + 2x$.

184 $\mathscr{B} = \{2-x, x^2-1, x^2-2, x^3-1, x^4\}$, $\mathbf{E} = \mathbf{P}_4$, $\vec{u} = x + x^4$.

185 $\mathscr{B} = \left\{ \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & -2 \\ -2 & 3 \end{bmatrix} \right\}$, $\mathbf{E}$ el espacio de matrices simétricas 2×2, $\vec{u} = \begin{bmatrix} 2 & 3 \\ 3 & 0 \end{bmatrix}$.

186 $\mathscr{B} = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \right\}$, $\mathbf{E} = \mathfrak{M}_2$, $\vec{u} = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}$.

En los ejercicios 187 a 196 encontrar la matriz cambio de base de la base canónica $\mathscr{B}_1$ a la base dada $\mathscr{B}_2$ del espacio indicado $\mathbf{E}$.

187 Del espacio $\mathbf{E}$ y la base $\mathscr{B}_2 = \mathscr{B}$ del ejercicio propuesto 171.

188 Del espacio $\mathbf{E}$ y la base $\mathscr{B}_2 = \mathscr{B}$ del ejercicio propuesto 172.

189 Del espacio $\mathbf{E}$ y la base $\mathscr{B}_2 = \mathscr{B}$ del ejercicio propuesto 177.

190 Del espacio $\mathbf{E}$ y la base $\mathscr{B}_2 = \mathscr{B}$ del ejercicio propuesto 178.

191 Del espacio $\mathbf{E}$ y la base $\mathscr{B}_2 = \mathscr{B}$ del ejercicio propuesto 179.

192 Del espacio $\mathbf{E}$ y la base $\mathscr{B}_2 = \mathscr{B}$ del ejercicio propuesto 180.

193 Del espacio $\mathbf{E}$ y la base $\mathscr{B}_2 = \mathscr{B}$ del ejercicio propuesto 181.

194 Del espacio $\mathbf{E}$ y la base $\mathscr{B}_2 = \mathscr{B}$ del ejercicio propuesto 182.

195 Del espacio $\mathbf{E}$ y la base $\mathscr{B}_2 = \mathscr{B}$ del ejercicio propuesto 184.

196 Del espacio $\mathbf{E}$ y la base $\mathscr{B}_2 = \mathscr{B}$ del ejercicio propuesto 185.

En los ejercicios 197 a 208 encontrar la matriz cambio de base *(i)* de la base $\mathscr{B}_2$ a la base $\mathscr{B}_1$ y *(ii)* de la base $\mathscr{B}_1$ a la base $\mathscr{B}_2$, en el espacio indicado $\mathbf{E}$.

197 $\mathscr{B}_1 = ((1,1),(1,0))$, $\mathscr{B}_2 = ((-2,1),(-3,2))$, $\mathbf{E} = \mathbb{R}^2$.

198 $\mathscr{B}_1 = ((2,1),(1,1))$, $\mathscr{B}_2 = ((4,1),(3,1))$, $\mathbf{E} = \mathbb{R}^2$.

199 $\mathscr{B}_1 = ((1,-4),(0,1))$, $\mathscr{B}_2 = ((5,2),(9,1))$, $\mathbf{E} = \mathbb{R}^2$.

200 $\mathscr{B}_1 = ((-3,5),(1,-2))$, $\mathscr{B}_2 = ((2,1),(1,0))$, $\mathbf{E} = \mathbb{R}^2$.

201 $\mathscr{B}_1 = ((2,1,4),(3,2,5),(0,-1,1))$, $\mathscr{B}_2 = ((1,2,-1),(0,-1,1),(0,0,-1))$, $\mathbf{E} = \mathbb{R}^3$.

202 $\mathscr{B}_1 = ((1,0,0),(-1,1,1),(1,1,2))$, $\mathscr{B}_2 = ((1,-2,-2),(-2,3,1),(-1,2,1))$, $\mathbf{E} = \mathbb{R}^3$.

203 $\mathscr{B}_1 = ((1,-1,3),(-1,2,-3),(2,-2,5))$,
$\mathscr{B}_2 = ((1,-3,2),(-1,-2,1),(1,0,0))$, $\mathbf{E} = \mathbb{R}^3$.

204 $\mathscr{B}_1 = ((1,-2,3,4),(-2,5,-6,-8),(1,-2,4,4),(3,-5,9,11))$, $\mathscr{B}_2$ la base canónica de $\mathbf{E} = \mathbb{R}^4$.

205 $\mathscr{B}_1 = (-x+x^2, 1-2x+2x^2, -2x+x^2)$, $\mathscr{B}_2 = (-3-3x-x^2, 5+2x+2x^2, x)$, $\mathbf{E} = \mathbf{P}_2$.

206 $\mathscr{B}_1 = (1-2x-2x^2, -2+3x+x^2, -1+2x+x^2)$, $\mathscr{B}_2 = (-2+3x-x^2, 5-x+x^2, 3-2x+x^2)$, $\mathbf{E} = \mathbf{P}_2$.

207 $\mathscr{B}_1 = (1-x+3x^2, -1+2x-3x^2, 2-2x+5x^2)$, $\mathscr{B}_2 = (1-x+3x^2, -1+2x-3x^2, 2-2x+5x^2)$, $\mathbf{E} = \mathbf{P}_2$.

208 $\mathscr{B}_1 = (-1+3x-2x^2+2x^3, 2-5x+4x^2-4x^3, -1+3x-3x^2-2x^3, 1-3x+2x^2-x^3)$, $\mathscr{B}_2 = (1,x,x^2,x^3)$, $\mathbf{E} = \mathbf{P}_3$.

209 Sean $\mathscr{B}_1 = (\vec{e}_1,\ldots,\vec{e}_n)$ y $\mathscr{B}_2 = (\vec{f}_1,\ldots,\vec{f}_n)$ bases ordenadas de $\mathbb{R}^n$, A la matriz que tiene por columnas a los vectores $\vec{e}_i$ y B la matriz que tiene por columnas a los vectores $\vec{f}_i$. Mostrar que si P es la matriz cambio de base de la base $\mathscr{B}_2$ a la base $\mathscr{B}_1$, entonces

$$P = A^{-1}B.$$

En los ejercicios 210 a 218 se define un operador $T:\mathbb{R}^n\to\mathbb{R}^n$. *(i)* Comprobar por simple inspección (cfr. el ejercicio propuesto 62*(c)* de este apartado) que el operador es lineal; *(ii)* encontrar la representación matricial $[T]_{\mathscr{B}}$ relativa a la base ordenada $\mathscr{B}$; *(iii)* calcular $[T(\vec{u})]_{\mathscr{B}}$ para el vector $\vec{u}$.

210 $T(x,y)=(x-y,x+3y)$, $\mathscr{B}=((-1,1),(2,1))$ en $\mathbb{R}^2$, $\vec{u}=(1,2)$.

211 $T(x,y)=(2x-y,x-y)$, $\mathscr{B}=((1,-1),(2,-1))$, en $\mathbb{R}^2$, $\vec{u}=(-1,3)$.

212 $T(x,y)=(2x+y,3x-2y)$, $\mathscr{B}=((3,2),(1,1))$, en $\mathbb{R}^2$, $\vec{u}=(2,-3)$.

213 $T(x,y)=(x,x+y)$, $\mathscr{B}=((-2,1),(1,-1))$, en $\mathbb{R}^2$, $\vec{u}=(1,0)$.

214 $T(x,y,z)=(x-y+z,x+3y,z)$, $\mathscr{B}=((2,1,4),(3,2,5),(0,-1,1))$ en $\mathbb{R}^3$, $\vec{u}=(-1,2,2)$.

215 $T(x,y,z)=(x-3y+2z,-x-y+z,x+y)$, $\mathscr{B}=((1,0,0),(-1,1,1),(1,1,2))$ en $\mathbb{R}^3$, $\vec{u}=(-1,0,1)$.

216 $T(x,y,z)=(x,y+z,x-y-z)$, $\mathscr{B}=((1,-1,3),(-1,2,-3),(2,-2,5))$ en $\mathbb{R}^3$, $\vec{u}=(2,1,0)$.

217 $T(x,y,z)=(z-y,x+3z,z)$, $\mathscr{B}=((0,-2,3),(1,-3,5),(0,3,4))$ en $\mathbb{R}^3$, $\vec{u}=(1,1,1)$.

218 $T(x_1,x_2,x_3,x_4)=(x_1,x_1-x_3,x_2-x_4,x_4)$,
$\mathscr{B}=((1,-2,3,4),(-2,5,-6,-8),(1,-2,4,4),(3,-5,9,11))$ en $\mathbb{R}^4$, $\vec{u}=(-1,0,1,1)$.

219 Hallar una transformación lineal $T:\mathbb{R}^2\to\mathbb{R}^2$ tal que $T(\hat{\imath})=\hat{\jmath}-\hat{\imath}$ y $T(\hat{\jmath})=2\hat{\imath}+\hat{\jmath}$ y para esta transformación:

- *(a)* Calcular $T(2\hat{\imath}-3\hat{\jmath})$ y $T^2(2\hat{\imath}-3\hat{\jmath})$.
- *(b)* Encontrar $T(x\hat{\imath}+y\hat{\jmath})$ para todo $x\hat{\imath}+y\hat{\jmath}\in\mathbb{R}^2$.
- *(c)* Encontrar $\operatorname{Ker}(T)$, $T(\mathbf{E})$, bases y dimensiones de sendos subespacios.
- *(d)* Hallar $[T]_{\mathscr{B}}$ la representación matricial de T relativa a la base canónica $\mathscr{B}=\{\hat{\imath},\hat{\jmath}\}$.
- *(e)* Hallar $[T]_{\mathscr{B}}$ la representación matricial de T relativa a la base $\mathscr{B}=\{\hat{\imath}-\hat{\jmath},3\hat{\jmath}+\hat{\imath}\}$.
- *(f)* Hallar $[T^2]_{\mathscr{B}}$ la representación matricial de T relativa a la base canónica $\mathscr{B}=\{\hat{\imath},\hat{\jmath}\}$.
- *(g)* Hallar $[T^2]_{\mathscr{B}}$ la representación matricial de T^2 relativa a la base $\mathscr{B}=\{\hat{\imath}-\hat{\jmath},3\hat{\jmath}+\hat{\imath}\}$.

220 Sea $\mathbf{P}_3$ el espacio de polinomios de grado a lo más 3.

- *(a)* Encontrar un operador lineal $T:\mathbf{P}_3\to\mathbf{P}_3$ tal que $T(1)=x-1$, $T(x)=x^2+2$, $T(x^2)=x-3$ y $T(x^3)=5$.
- *(b)* Calcular $T(1-2x+x^2-3x^3)$.
- *(c)* Encontrar $T(a_0+a_1x+a_2x^2+a_3x^3)$ para todo $a_0+a_1x+a_2x^2+a_3x^3\in\mathbf{P}_3$.
- *(d)* Encontrar $\operatorname{Ker}(T)$ y $T(\mathbf{P}_3)$, bases y dimensiones de sendos subespacios.
- *(e)* Hallar $[T]_{\mathscr{B}}$ la representación matricial de T relativa a la base canónica $\mathscr{B}=\{1,x,x^2,x^3\}$.
- *(f)* Hallar $[T]_{\mathscr{B}}$ la representación matricial de T^2 relativa a la base canónica $\mathscr{B}=\{1,x,x^2,x^3\}$.
- *(g)* Hallar $[T]_{\mathscr{B}}$ la representación matricial de T relativa a la base $\mathscr{B}=\{x-1,2-x,x^2-2,x^3+1\}$.
- *(h)* Hallar $[T]_{\mathscr{B}}$ la representación matricial de T^2 relativa a la base $\mathscr{B}=\{x-1,2-x,x^2-2,x^3+1\}$.

En los ejercicios 221 a 231 se define un operador $T:\mathbf{E}\to\mathbf{E}$ en el espacio $\mathbf{E}$ indicado. *(i)* Demostrar que T es lineal; *(ii)* encontrar la representación matricial $[T]_{\mathscr{B}}$ relativa a la base ordenada $\mathscr{B}$.

221 $T:\mathfrak{M}_2\to\mathfrak{M}_2$, $T(A)=A^t$, $\mathscr{B}=\left\{\begin{bmatrix}-1&1\\0&1\end{bmatrix},\begin{bmatrix}0&-1\\1&1\end{bmatrix},\begin{bmatrix}1&0\\-1&1\end{bmatrix},\begin{bmatrix}1&-1\\1&0\end{bmatrix}\right\}$

222 $T:\mathbf{P}_2\to\mathbf{P}_2$, $T(p)=p'$ (el operador derivación), $\mathscr{B}=\{x-1,x-2,x^2\}$.

223 $T:\mathbf{P}_3\to\mathbf{P}_3$, $T(p)=\int_0^x p'(t)dt$, $\mathscr{B}=\{1,x,x^2,x^3\}$.

224 $T:\mathbf{P}_3\to\mathbf{P}_3$, $T(p)=\int_0^x p'(t)dt$, $\mathscr{B}=\{1,x+1,x^2-1,x^3\}$.

225 $T:\mathbf{P}_3\to\mathbf{P}_3$, $T(p)=q$, donde $q(x)=p(x+1)$ $\mathscr{B}=(1,x,x^2,x^3)$.

226 $T:\mathbf{P}_2\to\mathbf{P}_2$, $T(p)=q$, donde $q(x)=p(x+1)$ $\mathscr{B}=(x-1,x-2,x^2)$.

227 $T:\mathbf{P}_n\to\mathbf{P}_n$, $T(a_0+a_1x+\cdots+a_nx^n)=\sum_{i=0}^n a_i+(\sum_{i=1}^n a_i)x+(\sum_{i=2}^n a_i)x^2+\cdots+a_nx^n$, $\mathscr{B}=(1,x,x^2,\ldots,x^n)$.

228 $T:\mathbf{P}_3\to\mathbf{P}_3$, $T(p)=q$ donde $q(x)=xp'(x)$, $\mathscr{B}=(1,x,x^2,x^3)$.

229 $T:\mathbf{P}_3\to\mathbf{P}_3$, $T(p)=q$ donde $q(x)=xp'(x)$, $\mathscr{B}=(1-x,x+1,x^2-1,x^3)$.

230 $T:\mathbf{E}\to\mathbf{E}$, $T(f)=f'$, donde $\mathbf{E}=\mathrm{gn}(e^x,e^{2x},e^{3x})$, $\mathscr{B}=(e^x+e^{2x},e^{2x}+e^{3x},e^x+e^{3x})$.

231 $T:\mathbf{P}_3\to\mathbf{P}_3$ el operador lineal definido por $T(1)=1+x$, $T(x)=(1+x)^2$, $T(x^2)=(1+x)^3$, $T(x^3)=x$.

232 Sean $\mathbf{E}$ un espacio vectorial de dimensión $n>1$, T un operador lineal nilpotente con índice de nilpotencia igual a n y $\vec{u}\in\mathbf{E}-\{\vec{0}_{\mathbf{E}}\}$ tal que $T^{n-1}(\vec{u})\neq\vec{0}_{\mathbf{E}}$. Por el ejercicio propuesto 172*(b)* de este capítulo, $\mathscr{B}=\{\vec{u},T(\vec{u}),\ldots,T^{n-1}(\vec{u})\}$ es una base de $\mathbf{E}$. Encontrar la representación matricial $[T]_{\mathscr{B}}$ de este operador relativa a la base $\mathscr{B}$.

En los ejercicios 233 a 237 $\mathbf{E}$ es un subespacio de $\mathscr{F}(\mathbb{R})$, el espacio de funciones, $\mathscr{B}$ es una base ordenada de $\mathbf{E}$ y $D:\mathbf{E}\to\mathbf{E}$ es el operador derivación, $D(f)=f'$. Hallar $[T]_{\mathscr{B}}$, la representación matricial de T relativa a la base $\mathscr{B}$.

233 $\mathscr{B}=(\cos x,\operatorname{sen} x)$.

234 $\mathscr{B}=(1,x,x^2)$.

235 $\mathscr{B}=(e^x,xe^x)$.

236 $\mathscr{B}=(e^{3x}\cos x,e^{3x}\operatorname{sen} x)$.

237 $\mathscr{B}=(\cos x,\operatorname{sen} x,x\cos x,x\operatorname{sen} x)$.

238 Sean $\mathbf{E}$ un espacio vectoriales de dimensión finita y $T_1,T_2\in\mathscr{L}(\mathbf{E},\mathbf{E})$. Probar que si $\mathscr{B}$ es una base de $\mathbf{E}$ y $[T_1]_{\mathscr{B}}$, $[T_2]_{\mathscr{B}}$ son las representaciones matriciales de T_1, T_2, respectivamente, relativas a esta base, entonces

$$\begin{aligned}[T_1+T_2]_{\mathscr{B}} &= [T_1]_{\mathscr{B}}+[T_2]_{\mathscr{B}} \quad \text{y}\\ [\alpha T_1]_{\mathscr{B}} &= \alpha[T_1]_{\mathscr{B}}.\end{aligned}$$

En los ejercicios 239 a 245 D es el operador derivación en $\mathbf{P}_3$ y $T:\mathbf{P}_3\to\mathbf{P}_3$ es el operador lineal $T(p)=q$, donde $q(x)=xp'(x)$ y $\mathscr{B}=(1,x,x^2,x^2)$ es la base ordenada canónica de $\mathbf{P}_3$. Hallar.

239 $[T]_{\mathscr{B}}$

240 $[T]_{\mathscr{B}}$

241 $[DT]_{\mathscr{B}}$

242 $[TD]_{\mathscr{B}}$

243 $[TD-DT]_{\mathscr{B}}$

244 $[T^2]_{\mathscr{B}}$

245 $[T^2D^2-D^2T^2]_{\mathscr{B}}$

En los ejercicios 246 a 255 encontrar la representación matricial del operador lineal $T:\mathbb{R}^n\to\mathbb{R}^n$ relativa a la base $\mathscr{B}_2$ procediendo de la manera siguiente:

(i) Hallar la matriz cambio de base, P, de la base $\mathscr{B}_1$ a la base canónica $\mathscr{B}$ de $\mathbb{R}^n$.

(ii) Utilizar la fórmula del teorema 5.15; esto es, $[T]_{\mathscr{B}_1}=P^{-1}[T]_{\mathscr{B}}P$.

246 $T:\mathbb{R}^2\to\mathbb{R}^2$, $T(x,y)=(2x-y,x+y)$, $\mathscr{B}_1=((1,-1),(2,-1))$.

247 $T:\mathbb{R}^2\to\mathbb{R}^2$, $T(x,y)=(x+3y,x-y)$, $\mathscr{B}_1=((2,3),(1,2))$.

248 $T:\mathbb{R}^2\to\mathbb{R}^2$, $T(x,y)=(y,x)$, $\mathscr{B}_1=((1,1),(2,1))$.

249 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(x,y,z)=(x-y+z,x+y+3z,-x-2y)$, $\mathscr{B}_1=((1,0,0),(-1,1,1),(1,1,2))$.

250 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(x,y,z)=(x-2y+z,3x-2y+4z,-x+2y+z)$, $\mathscr{B}_1=((1,2,-1),(-1,-1,1),(1,4,-2))$.

251 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(x,y,z)=(3x-2y+z,x+y,x-y)$, $\mathscr{B}_1=((1,0,0),(2,2,-1),(1,-1,1))$.

252 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(x,y,z)=(x-y+2z,x+y+x,x-2y-3z)$, $\mathscr{B}_1=((1,2,-1),(1,1,-1),(1,2,-2))$.

253 $T:\mathbb{R}^4\to\mathbb{R}^4$, $T(x_1,x_2,x_3,x_4)=(x_1-x_4,x_1+x_3,x_3-x_4,x_4)$,

$$\mathscr{B}_1=((1,2,1,3),(-1,-1,-1,-3),(1,2,2,3),(1,2,1,2)).$$

254 $T:\mathbb{R}^4\to\mathbb{R}^4$, $T(x_1,x_2,x_3,x_4)=(x_1-x_3+x_4,2x_1-x_2,x_3,x_4)$,

$$\mathscr{B}_1=((1,4,-2,-1),(1,3,-2,-1),(-1,-4,3,1),(1,4,-2,-2)).$$

255 $T:\mathbb{R}^4\to\mathbb{R}^4$, $T(x_1,x_2,x_3,x_4)=(x_1,x_1+x_2,x_1+x_2+x_3,x_1+x_2+x_3+x_4)$,

$$\mathscr{B}_1=((1,2,-1,-1),(-1,-1,1,1),(1,2,0,-1),(-1,-2,1,2)).$$

256 Sean $T:\mathbf{P}_3\to\mathbf{P}_3$ el operador lineal $T(p)=q$ donde $q(x)=xp'(x)$ y la base ordenada $\mathscr{B}_1=(2-x,x+1,x^2-1,x^3-3))$.

(*a*) Hallar la matriz cambio de base de la base $\mathscr{B}_1$ a la base canónica $\mathscr{B}=\{1,x,x^2,x^3\}$.

(*b*) Encontrar la representación matricial $[T]_{\mathscr{B}_1}$ del operador T relativa a la base $\mathscr{B}_1$.

257 Sean $T:\mathbf{P}_3\to\mathbf{P}_3$ el operador lineal $T(p)=p'$ y la base ordenada $\mathscr{B}_1=(2-x,x+1,x^2-1,x^3-3))$.

(*a*) Hallar la matriz cambio de base de la base $\mathscr{B}_1$ a la base canónica $\mathscr{B}=\{1,x,x^2,x^3\}$.

(*b*) Encontrar la representación matricial $[T]_{\mathscr{B}_1}$ del operador T relativa a la base $\mathscr{B}_1$.

258 Sean $T:\mathfrak{M}_2\to\mathfrak{M}_2$ el operador lineal $T(A)=A^t$ y la base ordenada

$$\mathscr{B}_1=\left(\begin{bmatrix}1&1\\0&0\end{bmatrix},\begin{bmatrix}1&2\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&1\end{bmatrix},\begin{bmatrix}0&1\\1&0\end{bmatrix}\right).$$

(*a*) Hallar la matriz cambio de base de la base $\mathscr{B}_1$ a la base canónica $\mathscr{B}$ de $\mathfrak{M}_2$.

(*b*) Encontrar la representación matricial $[T]_{\mathscr{B}_1}$ del operador T relativa a la base $\mathscr{B}_1$.

En los ejercicios 259 a 272 se define un operador $T:\mathbf{E}\to\mathbf{E}$. *(i)* Mostrar que T es lineal; *(ii)* calcular $\det(T)$, el determinante del operador; *(iii)* determinar si T es invertible y de ser así encontrar $T^{-1}(u)$ para todo $u\in\mathbf{E}$.

259 $T:\mathbb{R}^2\to\mathbb{R}^2$, $T(x,y)=(x-2y,x+y)$.

260 $T:\mathbb{R}^2\to\mathbb{R}^2$, $T(x,y)=(x+2y,2x+3y)$.

261 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(x,y)=(x+y+z,-2x+y+3z,3y+5z)$.

262 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(x,y,z)=(x-y+z,3x-2y+5z,2x-2y+z)$.

263 $T:\mathbb{R}^3\to\mathbb{R}^3$, $T(x,y,z)=(x-2y+3z,-x+4y,x+y+8z)$.

264 $T:\mathbb{R}^4\to\mathbb{R}^4$, $T(x_1,x_2,x_3,x_4)=(x_1-x_2+x_3+x_4,-x_1+2x_2+3x_3-4x_4,2x_3+x_4,5x_3+3x_4)$.

265 $T:\mathbb{R}^4\to\mathbb{R}^4$,

$$T(x_1,x_2,x_3,x_4)=(2x_1-x_2+3x_3+x_4,-x_1+2x_2+x_3+x_4,-x_1+5x_2+6x_3+4x_4,-x_1+x_2-x_3+2x_4).$$

266 $T:\mathfrak{M}_2\to\mathfrak{M}_2$, $T(A)=A^t$.

267 $T:\mathfrak{M}_2\to\mathfrak{M}_2$, $T(A)=\frac{1}{2}(A+A^t)$.

268 $T:\mathfrak{M}_2\to\mathfrak{M}_2$, $T(A)=\frac{1}{2}(A-A^t)$.

269 $T:\mathbf{P}_3\to\mathbf{P}_3$, $T(p)=p'$.

270 $T:\mathbf{P}_3\to\mathbf{P}_3$, $T(p)=q$ donde $q(x)=xp'(x)$.

271 $T:\mathbf{P}_3\to\mathbf{P}_3$, $T(p)=q$ donde $q(x)=p(x+1)$.

272 $T:\mathbf{P}_3\to\mathbf{P}_3$, $T(p)=q$ donde, si $p(x)=a_0+a_1x+a_2x^2+a_3x^3$, $p(x)=a_0+\frac{1}{2}a_1x+\frac{1}{3}a_2x^2+\frac{1}{4}a_3x^3$.

273 Sea $\mathfrak{M}_n$ el espacio vectorial de las matrices de orden n. Se recuerda que la traza de una matriz cuadrada $A=[a_{ij}]$ se define por $\text{tra}(A)=\sum_{i=1}^n a_{ii}$.

(a) Demostrar que si $A,B\in\mathfrak{M}_n$, entonces $\text{tra}(AB)=\text{tra}(BA)$.

(b) Probar que $\text{tra}(A)=\text{tra}(A^t)$ para toda matriz $A\in\mathfrak{M}_n$.

274 Sean $\mathbf{E}$ un espacio vectorial de dimensión finita y T un operador lineal en él, $\mathscr{B}_1$ y $\mathscr{B}_2$ un par de bases en él. Sean A y B las representaciones matriciales de T relativas a sendas bases, demostrar que

$$\text{tra}(A)=\text{tra}(B).$$

Es decir, cualquier representación matricial de T tiene la misma traza independientemente de la base que se elija en $\mathbf{E}$. Se define entonces la traza del operador T, $\text{tra}(T)$, como la traza de cualquier representación matricial de este operador. Luego $\text{tra}(T)$ es un invariante de T.

275 Sean $\mathbf{E}$ un espacio vectorial y T_1,T_2 un par de operadores lineales, mostrar que $\text{tra}(T_2T_1-T_1T_2)=0$.

276 Sean $\mathbf{E}$ y $\mathbf{F}$ espacios vectoriales de dimensiones finitas y $T_1,T_2\in\mathscr{L}(\mathbf{E},\mathbf{F})$. Probar que si $\mathscr{B}_1$ y $\mathscr{B}_2$ son bases de sendos espacios y $[T_1]_{\mathscr{B}_1}^{\mathscr{B}_2}$, $[T_2]_{\mathscr{B}_1}^{\mathscr{B}_2}$ son las representaciones matriciales de T_1, T_2, respectivamente, relativas a estas bases, entonces

$$\begin{aligned}[T_1+T_2]_{\mathscr{B}_1}^{\mathscr{B}_2} &= [T_1]_{\mathscr{B}_1}^{\mathscr{B}_2}+[T_2]_{\mathscr{B}_1}^{\mathscr{B}_2} \quad \text{y}\\ [\alpha T_1]_{\mathscr{B}_1}^{\mathscr{B}_2} &= \alpha[T_1]_{\mathscr{B}_1}^{\mathscr{B}_2}.\end{aligned}$$

277 Sean $\mathbf{E}$, $\mathbf{F}$ y $\mathbf{G}$ espacios vectoriales; $T_1:\mathbf{E}\to\mathbf{F}$ y $T_2:\mathbf{F}\to\mathbf{G}$ un par de funciones. Se define la operación composición de T_2 con T_1 como la función $T_2\circ T_1:\mathbf{E}\to\mathbf{G}$ dada por (cfr. el ejercicio resuelto 12 de esta sección)

$$(T_2\circ T_1)(\vec{u})=T_2(T_1(\vec{u}))\quad\forall\vec{u}\in\mathbf{E}.$$

Nuevamente se acostumbra denotar $T_2\circ T_1$ por T_2T_1.

(a) Probar que si T_1 y T_2 son lineales, entonces $T_2\circ T_1$ $(=T_2T_1)$ es también lineal.

(b) Sean $T_1,T_2:\mathbf{E}\to\mathbf{F}$ dos funciones y $T_3:\mathbf{F}\to\mathbf{G}$ una transformación lineal, mostrar que:

(i) $T_3\circ(T_1+T_2)=(T_3\circ T_1)+(T_3\circ T_2)$.

(ii) $T_3\circ(\alpha T_1)=\alpha(T_3\circ T_1)$, para todo $\alpha\in\mathbb{R}$.

(c) Sean $T_1,T_2:\mathbf{F}\to\mathbf{G}$ y $T_3:\mathbf{E}\to\mathbf{F}$ funciones, probar que

$$(T_1+T_2)\circ T_3=(T_1\circ T_3)+(T_2\circ T_3).$$

278 Sean **E**, **F** y **G** espacios vectoriales; $T_1 : \mathbf{E} \to \mathbf{F}$ y $T_2 : \mathbf{F} \to \mathbf{G}$ transformaciones lineales. Demostrar:

(a) Si T_2T_1 $(= T_2 \circ T_1)$ es inyectiva, entonces T_1 es inyectiva.

(b) Si T_2T_1 es suprayectiva, entonces T_2 es suprayectiva.

(c) Si T_1 y T_2 son biyectivas, entonces T_2T_1 es biyectiva.

279 Sean **E**, **F** y **G** espacios vectoriales de dimensiones finitas con sendas bases $\mathscr{B}_1$, $\mathscr{B}_2$ y $\mathscr{B}_3$. Si $T_1 : \mathbf{E} \to \mathbf{F}$ y $T_2 : \mathbf{F} \to \mathbf{G}$ son transformaciones lineales, demostrar que la representación matricial de $T_2T_1 = T_2 \circ T_1$ está dada por

$$[T_2T_1]_{\mathscr{B}_1}^{\mathscr{B}_3} = [T_2]_{\mathscr{B}_2}^{\mathscr{B}_3}[T_1]_{\mathscr{B}_1}^{\mathscr{B}_2}.$$

En los ejercicios 280 a 295 **E**, **F** son espacios vectoriales con sendas bases $\mathscr{B}_1$ y $\mathscr{B}_2$ y $T : \mathbf{E} \to \mathbf{F}$ es una transformación definida por la fórmula dada.

(i) Comprobar que la transformación es lineal.

(ii) Verificar que efectivamente $\mathscr{B}_1$ y $\mathscr{B}_2$ son bases de **E** y **F**, respectivamente.

(iii) Encontrar la representación matricial de T relativa a las bases $\mathscr{B}_1$ y $\mathscr{B}_2$; es decir, $[T]_{\mathscr{B}_1}^{\mathscr{B}_2}$.

280 $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(x,y,z) = (x-y+z, x+2y-z)$,
$\mathscr{B}_1 = ((1,-1,1),(1,0,-1),(-1,-1,1))$, $\mathscr{B}_2 = ((1,3),(1,2))$.

281 $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(x,y,z) = (x+y-z, x+y-3z)$,
$\mathscr{B}_1 = ((0,-1,1),(1,0,-1),(1,-1,1))$, $\mathscr{B}_2 = ((2,1),(1,1))$.

282 $T : \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (x-y, x+2y, x+y)$, $\mathscr{B}_1 = ((1,1),((-1,1))$, $\mathscr{B}_2 = ((1,0,0),(-1,1,1),(1,1,2))$.

283 $T : \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (x, x+y, x-y)$, $\mathscr{B}_1 = ((2,1),((-1,0))$,
$\mathscr{B}_2 = ((1,2,-1),(-1,-1,1),(1,4,-2))$.

284 $T : \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (x-2y, 2x+y, x-3y)$, $\mathscr{B}_1 = ((1,-1),(1,0))$,
$\mathscr{B}_2 = ((1,2,-1),(1,1,-1),(1,2,-2))$.

285 $T : \mathbb{R}^4 \to \mathbb{R}^3$, $T(x_1,x_2,x_3,x_4) = (x_1-x_4, x_1+x_3, x_2+x_3-x_4)$,
$\mathscr{B}_1 = ((1,-1,1,1),(1,0,-1,1),(0,1,1,1),(0,1,0,-1))$,
$\mathscr{B}_2 = ((1,5,2),(-1,-4,-2),(1,5,1))$.

286 $T : \mathbb{R}^4 \to \mathbb{R}^3$, $T(x_1,x_2,x_3,x_4) = (x_1-x_2, x_1-2x_3, x_2-x_3-x_4)$,
$\mathscr{B}_1 = ((1,0,-1,1),(0,-1,1,1),(-1,0,0,1),(0,0,0,1))$,
$\mathscr{B}_2 = ((1,6,-1),(1,5,-1),(-1,-6,2))$.

287 $T : \mathbb{R}^2 \to \mathbb{R}^4$, $T(x,y) = (x+2y, x-y, x+y, 3x-y)$, $\mathscr{B}_1 = ((1,-2),(-1,1))$,
$\mathscr{B}_2 = ((1,2,1,3),(-1,-1,-1,-3),(1,2,2,3),(1,2,1,2))$.

288 $T : \mathbb{R}^2 \to \mathbb{R}^4$, $T(x,y) = (2x-3y, x+4y, x-y, 3x-2y)$, $\mathscr{B}_1 = ((1,1),(1,-1))$,
$\mathscr{B}_2 = ((1,4,-2,-1),(1,3,-2,-1),(-1,-4,3,1),(1,4,-2,-2))$.

289 $T:\mathbb{R}^3\to\mathbb{R}^4$, $T(x,y,z)=(x-y+z,x+z,x+y-z,x+y+z)$, $\mathscr{B}_1=((1,0,1),(0,1,1),(1,1,0))$, $\mathscr{B}_2=((1,2,-1,-1),(-1,-1,1,1),(1,2,0,-1),(-1,-2,1,2))$.

290 $T:\mathfrak{M}_2\to\mathbf{P}_2$, $T\left(\begin{bmatrix} a & b\\ c & d\end{bmatrix}\right)=a+b++2dx+bx^2$,
$\mathscr{B}_1=\left(\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)$, $\mathscr{B}_2=(1,x,x^2)$.

291 $T:\mathfrak{M}_2\to\mathbf{P}_2$, $T\left(\begin{bmatrix} a & b\\ c & d\end{bmatrix}\right)=a+b+c+2dx+bx^2$,
$\mathscr{B}_1=\left(\begin{bmatrix}1&1\\0&0\end{bmatrix},\begin{bmatrix}1&2\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&1\end{bmatrix},\begin{bmatrix}0&1\\1&0\end{bmatrix}\right)$, $\mathscr{B}_2=(1-x,x-2,x^2)$.

292 $T:\mathbf{P}_2\to\mathfrak{M}_2$, $T(p)=\begin{bmatrix}p'(0) & 3p(1)\\ 0 & p''(2)\end{bmatrix}$,
$\mathscr{B}_1=\left(\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)$, $\mathscr{B}_2=(1,x,x^2)$.

293 $T:\mathbf{P}_2\to\mathfrak{M}_2$, $T(p)=\begin{bmatrix}p'(0) & 3p(1)\\ 0 & p''(2)\end{bmatrix}$,
$\mathscr{B}_1=\left(\begin{bmatrix}1&1\\0&0\end{bmatrix},\begin{bmatrix}1&2\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&1\end{bmatrix},\begin{bmatrix}0&1\\1&0\end{bmatrix}\right)$, $\mathscr{B}_2=(1-x,x-2,x^2)$.

294 $T:\mathfrak{M}_2\to\mathbb{R}$, $T(A)=\text{tra}(A)$,
$\mathscr{B}_1=\left(\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)$, $\mathscr{B}_2=\{1\}$.

295 $T:\mathfrak{M}_2\to\mathbb{R}$, $T(A)=\text{tra}(A)$,
$\mathscr{B}_1=\left(\begin{bmatrix}1&1\\0&0\end{bmatrix},\begin{bmatrix}1&2\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&1\end{bmatrix},\begin{bmatrix}0&1\\1&0\end{bmatrix}\right)$, $\mathscr{B}_2=\{1\}$.

En los ejercicios 296 a 302 encontrar $[T(\vec{u})]_{\mathscr{B}_1}^{\mathscr{B}_2}$ si:

296 T, $\mathscr{B}_1$ y $\mathscr{B}_2$ son la transformación lineal y las bases del ejercicio propuesto 280 de esta sección y $\vec{u}=(-1,1,2)$.

297 T, $\mathscr{B}_1$ y $\mathscr{B}_2$ son la transformación lineal y las bases del ejercicio propuesto 283 de esta sección y $\vec{u}=(-1,4)$.

298 T, $\mathscr{B}_1$ y $\mathscr{B}_2$ son la transformación lineal y las bases del ejercicio propuesto 286 de esta sección y $\vec{u}=(-1,0,2,2)$.

299 T, $\mathscr{B}_1$ y $\mathscr{B}_2$ son la transformación lineal y las bases del ejercicio propuesto 289 de esta sección y $\vec{u}=(2,1,4)$.

300 T, $\mathscr{B}_1$ y $\mathscr{B}_2$ son la transformación lineal y las bases del ejercicio propuesto 291 de esta sección y $\vec{u}=\begin{bmatrix}-1 & 1\\ 1 & 2\end{bmatrix}$.

301 T, $\mathscr{B}_1$ y $\mathscr{B}_2$ son la transformación lineal y las bases del ejercicio propuesto 293 de esta sección y $p(x) = 1 - x + x^2$.

302 T, $\mathscr{B}_1$ y $\mathscr{B}_2$ son la transformación lineal y las bases del ejercicio propuesto 295 de esta sección y $\vec{u} = \begin{bmatrix} -1 & 1 \\ 1 & 3 \end{bmatrix}$.

En los ejercicios 303 a 306 encontrar la representación matricial del operador lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ relativa a las bases $\mathscr{B}_1'$, $\mathscr{B}_2'$ de $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente, procediendo de la manera siguiente:

(i) Hallar la matriz cambio de base, P, de la base $\mathscr{B}_1'$ a la base canónica $\mathscr{B}_1$ de $\mathbb{R}^n$.

(ii) Hallar la matriz cambio de base, Q, de la base $\mathscr{B}_2'$ a la base canónica $\mathscr{B}_2$ de $\mathbb{R}^m$.

(ii) Utilizando la fórmula el teorema 5.19; esto es, $[T]_{\mathscr{B}_1'}^{\mathscr{B}_2'} = Q^{-1}[T]_{\mathscr{B}_1}^{\mathscr{B}_2}P$.

303 $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(x,y,z) = (x - y + z, x + y - z)$,
$\mathscr{B}_1' = ((-1,0,1),(1,0,0),(0,-1,1))$, $\mathscr{B}_2' = ((2,1),(1,1))$.

304 $T : \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (x + 2y, -x + 3y, x - y)$, $\mathscr{B}_1' = ((1,1),(-1,1))$,
$\mathscr{B}_2' = ((1,5,2),(-1,-4,-2),(1,5,1))$.

305 $T : \mathbb{R}^4 \to \mathbb{R}^3$, $T(x_1,x_2,x_3,x_4) = (x_1, x_1 - x_3, x_2 + x_4)$,
$\mathscr{B}_1' = (1,0,-1,1),(0,1,1,1),(-1,1,1,0),(0,-1,0,0))$,
$\mathscr{B}_2' = ((1,6,1),(1,5,-1),(-1,-6,2))$.

306 $T : \mathbb{R}^4 \to \mathbb{R}^2$, $T(x_1,x_2,x_3,x_4) = (x_1 - x_2 + x_3 - x_4, x_1 + x_2 + x_3 + x_4)$,
$\mathscr{B}_1' = (1,0,-1,1),(0,1,1,1),(-1,1,1,0),(0,-1,0,0))$,
$\mathscr{B}_2' = ((3,1),(5,2))$.

307 Sea π un proyector (cfr. el ejercicio propuesto 150 de esta sección) en un espacio vectorial **E** que tiene dimensión finita n. Si π tiene rango r; es decir $\dim(\pi(\mathbf{E})) = r$, mostrar que **E** posee una base $\mathscr{B}$ tal que

$$[\pi]_{\mathscr{B}} = \begin{bmatrix} I_r & \mathscr{O}_1 \\ \mathscr{O}_2 & \mathscr{O}_3 \end{bmatrix}$$

donde I_r es la identidad $r \times r$ y las submatrices $\mathscr{O}_1$, $\mathscr{O}_2$ y $\mathscr{O}_3$ son las matrices cero de tamaños $r \times (n-r)$, $(n-r) \times r$ y $(n-r) \times (n-r)$, respectivamente.

308 Sean **E** un espacio vectorial de dimensión finita n, $S < \mathbf{E}$ y T un operador lineal en **E**. Se dice que S es T-invariante si $T(S) \subset S$. Mostrar que si S tiene dimensión r, entonces **E** posee una base $\mathscr{B}$ tal que

$$[T]_{\mathscr{B}} = \begin{bmatrix} A & B \\ \mathscr{O} & C \end{bmatrix}$$

donde A es una matriz cuadrada de orden r, $\mathscr{O}$ es la matriz cero de tamaño $(n-r) \times r$, $B \in \mathfrak{M}_{r \times (n-r)}$ y $C \in \mathfrak{M}_{(n-r) \times (n-r)}$.

309 Sean **E** un espacio vectorial de dimensión finita, $S_1 < \mathbf{E}$ y $T : \mathbf{E} \to \mathbf{E}$ una proyección sobre S_1 (cfr. el ejercicio resuelto 17 de esta sección). Mostrar que **E** posee una base $\mathscr{B}$ tal que $[T]_{\mathscr{B}}$ es diagonal.

310 Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ la transformación lineal que satisface $T(1,0) = (0,1,1)$ y $T(0,1) = (0,-1,1)$.

(a) Calcular $T(x,y)$ para todo $(x,y) \in \mathbb{R}^2$.

(b) Encontrar bases y dimensiones de $\text{Ker}(T)$ y $T(\mathbb{R}^2)$.

(c) Hallar la representación matricial de T para las bases canónicas de $\mathbb{R}^2$ y $\mathbb{R}^3$.

(d) Encontrar bases $\mathscr{B}_1 = \{\vec{e}_1, \vec{e}_2\}$ y $\mathscr{B}_2 = \{\vec{f}_1, \vec{f}_2, \vec{f}_3\}$ de $\mathbb{R}^2$ y $\mathbb{R}^3$, respectivamente, tal que $[T]_{\mathscr{B}_1}^{\mathscr{B}_2}$ sea diagonal (cfr. el teorema 5.20, pág. 450).

311 Resolver el ejercicio precedente para la transformación T, con $T(1,0) = (-1,0,1)$, $T(0,1) = (1,1,1)$.

312 Sean $D, T : \mathbf{P}_3 \to \mathbf{P}_3$ los operadores lineales en $D(p) = p'$, $T(p)(x) = xp'(x)$ y TD en $\mathbf{P}_3$. Si $\mathbf{F} = (TD)(\mathbf{E})$, entonces $TD : \mathbf{P}_3 \to \mathbf{F}$; encontrar bases $\{\vec{e}_i\}$ y $\{\vec{f}_i\}$ de $\mathbf{E}$ y $\mathbf{F}$, respectivamente, tal que $[TD]_{\{\vec{e}_i\}}^{\{\vec{f}_i\}}$ sea diagonal (cfr. el teorema 5.20).

313 Sea $f : \mathbb{R}^2 \to \mathbb{R}$ el funcional lineal definido por $f(1,2) = -1$ y $f(1,1) = 2$. Encontrar $f(x,y)$ para todo $(x,y) \in \mathbb{R}^2$. En particular, hallar $f(2,-3)$.

314 Sea $f : \mathbb{R}^2 \to \mathbb{R}$ el funcional lineal definido por $f(1,-1) = 3$ y $f(3,1) = -4$. Encontrar $f(x,y)$ para todo $(x,y) \in \mathbb{R}^2$. En particular, hallar $f(-2,1)$.

En los ejercicios 315 a 325 $\phi : \mathbf{E} \to \mathbb{R}$ es una función definida en el espacio $\mathbf{E}$. Determinar si $\phi \in \mathbf{E}^*$; es decir, si ϕ es lineal.

315 $\mathbf{E} = C[0,1]$, $\phi(f) = f(0)$.

316 $\mathbf{E} = C[a,b]$, $\phi(f) = \int_a^b f(x)dx$.

317 $\mathbf{E} = \mathbf{P}$, $\phi(p) = p(0)$.

318 $\mathbf{E} = \mathbf{P}$, $\phi(p) = p'(0)$.

319 $\mathbf{E} = \mathfrak{M}_n$, $\phi(A) = \text{tra}(A)$.

320 $\mathbf{E} = \mathbf{P}$, $\phi(p) = p(2)$.

321 $\mathbf{E} = \mathbf{P}_2$, $\phi(a_0 + a_1x + a_2x^2) = a_0 + 2a_1 - 3a_2$.

322 $\mathbf{E} = \mathfrak{M}_2$, $\phi(A) = \det(A)$.

323 Sean $\vec{u}_2, \ldots, \vec{u}_n$, $n-1$ vectores fijos de $\mathbb{R}^n$, $\mathbf{E} = \mathbb{R}^n$ y $\phi(\vec{u}) = \det(A)$ donde A es la matriz con primera fila $\vec{u}$, e i-ésima fila $\vec{u}_i$, $i = 2, \ldots, n$.

324 $\mathbf{E} = \mathbb{R}^n$, $\phi(\vec{x}) = \vec{x} \cdot \vec{a}$ donde $\vec{a}$ es un vector fijo de $\mathbb{R}^n$.

325 $\mathbf{E} = \mathbb{R}^2$, $\phi(x,y) = x^2 - y^2$.

En los ejercicios 326 a 333, para cada espacio $\mathbf{E}$, $\mathscr{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ es una base del mismo y $\mathbf{E}^*$ el espacio dual definido en el ejercicio resuelto 30 de esta sección. Utilizar ese ejercicio para construir la base dual $\mathscr{B}^*$ para el espacio $\mathbf{E}^*$ (cfr. el ejercicio resuelto 31 de este capítulo).

326 $\mathbf{E} = \mathbb{R}^2$, $\mathscr{B} = \{(1,1),(2,1)\}$.

327 $\mathbf{E} = \mathbb{R}^2$, $\mathscr{B} = \{(2,-3),(-1,1)\}$.

328 $\mathbf{E} = \mathbb{R}^3$, $\mathscr{B} = \{(1,2,3),(-1,-1,-3),(1,2,2)\}$.

329 $\mathbf{E} = \mathbb{R}^3$, $\mathscr{B} = \{(1,4,3),(0,1,2),(1,4,2)\}$.

330 $\mathbf{E} = \mathbf{P}_2$, $\mathscr{B} = \{1,x,x^2\}$.

331 $\mathbf{E} = \mathbf{P}_1$, $\mathscr{B} = \{1-x, 2+x\}$.

332 $\mathbf{E} = \mathbf{P}_3$, $\mathscr{B} = \{1-x, 2+x, x^2-1, x^3\}$.

333 $\mathbf{E} = \mathbb{R}^n$, $\mathscr{B} = \{\vec{e}_i\}$ la base canónica.

334 Sea $\mathbf{E}$ un espacio vectorial con bases $\mathscr{B}_1 = \{\vec{e}_1, \ldots, \vec{e}_n\}$ y $\mathscr{B}_2 = \{\vec{f}_1, \ldots, \vec{f}_m\}$; $\mathbf{E}^*$ el espacio dual definido en el ejercicio resuelto 30 de esta sección. Si $\mathscr{B}_1^*$ y $\mathscr{B}_2^*$ son las bases duales de $\mathscr{B}_1$ y $\mathscr{B}_2$, respectivamente y P es la matriz cambio de base de la base $\mathscr{B}_1$ a la base $\mathscr{B}_2$, mostrar que la matriz cambio de base de la base $\mathscr{B}_1^*$ a la base $\mathscr{B}_2^*$ es $(P^t)^{-1}$.

En los ejercicios 335 a 345:

(i) Probar que el conjunto $\mathscr{B}^*$ es una base del espacio dual $\mathbf{E}^*$ del espacio dado $\mathbf{E}$.

(ii) Encontrar una base $\mathscr{B}$ del espacio $\mathbf{E}$ cuya base dual es $\mathscr{B}^*$.

335 $\mathbf{E} = \mathbb{R}^n$, $\mathscr{B}^* = \{\pi_1, \ldots, \pi_n\}$, donde $\pi_i(x_1, \ldots, x_i, \ldots, x_n) = x_i$ es la i-ésima proyección del vector $(x_1, \ldots, x_i, \ldots, x_n)$.

336 $\mathbf{E} = \mathbb{R}^2$, $\mathscr{B}^* = \{\phi_1, \phi_2\}$ donde $\phi_1(x,y) = x - 2y$ y $\phi_2(x,y) = -x + y$.

337 $\mathbf{E} = \mathbb{R}^2$, $\mathscr{B}^* = \{\phi_1, \phi_2\}$ donde $\phi_1(x,y) = 3x + 4y$ y $\phi_2(x,y) = 2x + 3y$.

338 $\mathbf{E} = \mathbb{R}^3$, $\mathscr{B}^* = \{\phi_1, \phi_2, \phi_3\}$ donde $\phi_1(x,y,z) = x + 3y - 2z$, $\phi_2(x,y,z) = -x - 2y + 2z$, $\phi_3(x,y,z) = y - z$.

339 $\mathbf{E} = \mathbb{R}^3$, $\mathscr{B}^* = \{\phi_1, \phi_2, \phi_3\}$ donde $\phi_1(x,y,z) = x + y + 2z$, $\phi_2(x,y,z) = y + z$, $\phi_3(x,y,z) = -x - y - z$.

340 $\mathbf{E} = \mathbf{P}_1$, $\mathscr{B}^* = \{\phi_1, \phi_2\}$ donde $\phi_1(p) = \int_0^1 p(x)dx$ y $\phi_2(p) = \int_0^2 p(x)dx$, para todo $p \in \mathbf{P}_1$.

341 $\mathbf{E} = \mathbf{P}_1$, $\mathscr{B}^* = \{\phi_1, \phi_2\}$ donde $\phi_1(a+bx) = 4a + 3b$ y $\phi_2(a+bx) = 3a + 2b$, para todo $p = a + bx \in \mathbf{P}_1$.

342 $\mathbf{E} = \mathbf{P}_2$, $\mathscr{B}^* = \{\phi_1, \phi_2, \phi_3\}$ donde $\phi_1(p) = p(0)$, $\phi_2(p) = p'(0)$ y $\phi_3(p) = p''(0)$, para todo $p(x) = a_0 + a_1x + a_2x^2 \in \mathbf{P}_2$.

343 $\mathbf{E} = \mathbf{P}_2$, $\mathscr{B}^* = \{\phi_1, \phi_2, \phi_3\}$ donde $\phi_1(p) = \int_0^1 p(x)dx$, $\phi_2(p) = p(0)$ y $\phi_3(p) = p'(1)$, para todo $p(x) = a_0 + a_1x + a_2x^2 \in \mathbf{P}_2$.

344 $\mathbf{E} = \mathfrak{M}_2$, $\mathscr{B}^* = \{\phi_1, \phi_2, \phi_3, \phi_4\}$ donde $\phi_1(A) = a - b$, $\phi_2(A) = 2a - b$, $\phi_3(A) = b + c$ y $\phi_4(A) = -a + b + c - d$, para toda $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathfrak{M}_2$.

345 $\mathbf{E} = \mathfrak{M}_2$, $\mathscr{B}^* = \{\phi_1, \phi_2, \phi_3, \phi_4\}$ donde $\phi_1(A) = a - b + c - d$, $\phi_2(A) = 2a - b + 2c - 2d$, $\phi_3(A) = a - b + 2c - d$ y $\phi_4(A) = -3a + 3b - 3c + 2d$, para toda $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathfrak{M}_2$.

346 Si $\mathbf{E}$ es un espacio vectorial de dimensión finita n y $f \in \mathbf{E}^*$ y existe $\vec{u} \in \mathbf{E}$ tal que $f(\vec{u}) \neq 0$, ¿cuál es el valor de $\dim(\operatorname{Ker}(f))$?

347 Sean $\mathbf{E}$ un espacio vectorial y $f, g \in \mathbf{E}^*$ tales que $f(\vec{u}) = 0 \Rightarrow g(\vec{u}) = 0$; esto es, $\operatorname{Ker}(f) \subset \operatorname{Ker}(g)$. Probar que $g = kf$ para algún $k \in \mathbb{R}$.

348 Mostrar que si $\mathbf{E}$ es un espacio vectorial, $f, g \in \mathbf{E}^*$, $h : \mathbf{E} \to \mathbb{R}$ se define por $h(\vec{u}) = f(\vec{u})g(\vec{u})$ y $h \in \mathbf{E}^*$, entonces $f = \theta$ o $g = \theta$, el funcional constante cero.

En los ejercicios 349 a 356 $\mathbf{E}$ es un espacio vectorial con producto interior $\langle \cdot, \cdot \rangle$ y $f : \mathbf{E} \to \mathbb{R}$ es un funcional. Comprobar que f es lineal y consultar el ejercicio resuelto 37 de este capítulo para encontrar el elemento $\vec{u}_f$ en $\mathbf{E}$ tal que $f(\vec{x}) = \langle \vec{x}, \vec{u}_f \rangle$ para todo $\vec{x} \in \mathbf{E}$.

349 $\mathbf{E} = \mathbb{R}^2$, $f(x, y) = 2x - 3y$, $\langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y}$.

350 $\mathbf{E} = \mathbb{R}^2$, $f(x, y) = x - y$, $\langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y}$.

351 $\mathbf{E} = \mathbb{R}^3$, $f(x, y, z) = x - y + z$, $\langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y}$.

352 $\mathbf{E} = \mathbb{R}^3$, $f(x, y, z) = 2x + 3y - z$, $\langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y}$.

353 $\mathbf{E} = \mathfrak{M}_2$, $f(A) = a_{11} - 2a_{12} + 3a_{21} + a_{22}$ para toda $A = [a_{ij}] \in \mathfrak{M}_2$, $\langle A, B \rangle = \operatorname{tra}(B^t A)$.

354 $\mathbf{E} = \mathfrak{M}_2$, $f(A) = 2a_{11} - a_{12} + a_{21} - a_{22}$ para toda $A = [a_{ij}] \in \mathfrak{M}_2$, $\langle A, B \rangle = \operatorname{tra}(B^t A)$.

355 $\mathbf{E} = \mathbf{P}_2$, $f(p(x)) = \int_0^1 p(x)dx$ para todo $p \in \mathbf{P}_2$, $\langle p, q \rangle = \int_{-1}^1 p(x)q(x)dx$.

356 $\mathbf{E} = \mathbf{P}_2$, $f(p(x)) = p'(0)$ para todo $p \in \mathbf{P}_2$, $\langle p, q \rangle = \int_{-1}^1 p(x)q(x)dx$.

357 Si $\mathbf{E}$ es un espacio vectorial de dimensión finita dotado de un producto interior $\langle \cdot, \cdot \rangle$, por el ejercicio resuelto 37 de este capítulo, para cada $f \in \mathbf{E}^*$ existe un único $\vec{u}_f \in \mathbf{E}$ tal que $f(\vec{x}) = \langle \vec{x}, \vec{u}_f \rangle$ para todo $\vec{x} \in \mathbf{E}$. Probar que la aplicación $\Phi : \mathbf{E}^* \to \mathbf{E}$ definida por $\Phi(f) = \vec{u}_f$ es un isomorfismo.

358 Sea $\mathbf{E}$ un espacio vectorial con producto interior $\langle \cdot, \cdot \rangle$. Sean $\vec{u} \in \mathbf{E}$ y $f_{\vec{u}} : \mathbf{E} \to \mathbb{R}$ definida por $f_{\vec{u}}(\vec{x}) = \langle \vec{x}, \vec{u} \rangle$ para todo $\vec{x} \in \mathbf{E}$.

(a) Probar que $f_{\vec{u}} \in \mathbf{E}^*$.

(*b*) Si $\mathbf{E}$ tiene dimensión finita, demostrar que la función $\Psi : \mathbf{E} \to \mathbf{E}^*$ definida, para cada $\vec{u} \in \mathbf{E}$, por $\Psi(\vec{u}) = f_{\vec{u}}$, es un isomorfismo. ¿Qué función es la inversa de Ψ?

359 Sea $\mathbf{E}$ un espacio de dimensión finita. Si $S \subset \mathbf{E}$ es no vacío y $f \in \mathbf{E}^*$, se dice que f es *ortogonal* a S si $f(\vec{u}) = 0$ para todo $\vec{u} \in S$. Al conjunto de elementos en $\mathbf{E}$ que son ortogonales a S se le denota por $S^{\perp}$.

(*a*) Mostrar que $S^{\perp}$ es un subespacio de $\mathbf{E}^*$.

(*b*) Probar que si S es un subespacio de $\mathbf{E}$, entonces

$$\dim(S) + \dim(S^{\perp}) = \dim(\mathbf{E}).$$

En los ejercicios 360 a 364 hallar el operador adjunto T^*, definido en el ejercicio propuesto 38 de este capítulo, para el operador lineal T en el espacio $\mathbf{E}$ dotado del producto interior $\langle \cdot, \cdot \rangle$.

360 $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x+y, 2x-5y)$, $\langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y}$.

361 $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (x+3y, 2x+z, -4y)$, $\langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y}$.

362 $T : \mathbb{R}^4 \to \mathbb{R}^4$, $T(x_1,x_2,x_3,x_4)) = (x_1 - x_2, 2x_3 + x_4, x_1 - 2x_2, x_1 + 3x_4)$, $\langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y}$.

363 $T : \mathbf{P}_3 \to \mathbf{P}_3$, $T(p) = p'$, $\langle p, q \rangle = \int_0^1 p(x)q(x)dx$.

364 $T : \mathbf{P}_2 \to \mathbf{P}_2$, $T(p) = q$, donde $q(x) = xp'(x)$, $\langle p, q \rangle = \int_0^1 p(x)q(x)dx$.

365 Probar que para el operador lineal $D : \mathbf{P} \to \mathbf{P}$ definido por $D(p) = p'$ no existe el operador adjunto D^* si se considera el producto interior en $\mathbf{P}$: $\langle p, q \rangle = \int_0^1 p(x)q(x)dx$.

366 Sean $\mathbf{E}$ un espacio de dimensión finita con producto interior $\langle \cdot, \cdot \rangle$, $T : \mathbf{E} \to \mathbf{E}$ un operador lineal y T^* el operador adjunto correspondiente definido en el ejercicio resuelto 38 de este capítulo. Mostrar que si S es un subespacio T-invariante, entonces $S^{\perp}$ es T^*-invariante (cfr. el ejercicio propuesto 308).

367 (**Transpuesta de una transformación lineal**). Sean $\mathbf{E}$, $\mathbf{F}$ dos espacios vectoriales y $T : \mathbf{E} \to \mathbf{F}$ una transformación lineal. Se define $T^t : \mathbf{F}^* \to \mathbf{E}^*$, para cada $\phi \in \mathbf{F}^*$, por

$$T^t(\phi) = \phi \circ T,$$

la composición de las transformaciones ϕ y T (cfr. el ejercicio propuesto 277). A T^t se le llama la *transpuesta* de la transformación T.

(*a*) Probar que $T^t \in \mathbf{E}^*$; es decir, que T es lineal.

(*b*) Si $\mathbf{E}$ y $\mathbf{F}$ tienen dimensiones finitas, $\mathscr{B}_1$, $\mathscr{B}_2$ son bases de sendos espacios y $A = [T]_{\mathscr{B}_1}^{\mathscr{B}_2}$ es la representación de T relativa a las bases $\mathscr{B}_1$, $\mathscr{B}_2$, demostrar que

$$[T^t]_{\mathscr{B}_1^*}^{\mathscr{B}_2^*} = A^t,$$

donde $\mathscr{B}_1^*$ y $\mathscr{B}_2^*$ son las bases duales de $\mathscr{B}_1$ y $\mathscr{B}_2$, respectivamente.

(*c*) Si $\mathbf{E}$ y $\mathbf{F}$ tienen dimensiones finitas, probar que

$$\mathrm{Rang}(T) = \mathrm{Rang}(T^t).$$

En los ejercicios 368 a 376 $\mathbf{E}$ es un espacio vectorial de dimensión finita y para cualquier $S \subset \mathbf{E}$, S^0 denota el subespacio anulador de S definido en el ejercicio resuelto 40 de este capítulo.

368 Probar que $S^0 = (\mathscr{L}(S))^0$, donde $\mathscr{L}(S)$ es el espacio generado por S (cfr. el ejercicio resuelto 31 del capítulo 3).

369 Si W es un subespacio de $\mathbf{E}$ y $\vec{u} \in \mathbf{E} - W$, probar que existe $f \in W^0$ tal que $f(\vec{u}) \neq 0$.

370 Probar que $S^{00} = \mathscr{L}(\Phi(S))$; es decir, $(S^0)^0 = \mathscr{L}(\Phi(S))$ donde Φ es el isomorfismo entre los espacios $\mathbf{E}$ y $\mathbf{E}^{**}$ definido en el ejercicio resuelto 32 de este capítulo y $\mathscr{L}(\Phi(S))$ es el subespacio generado por el conjunto $\Phi(S)$ (cfr. el ejercicio resuelto 31 del capítulo 3).

371 Mostrar que si W_1, W_2 son dos subespacios de $\mathbf{E}$, entonces $W_1 = W_2 \Leftrightarrow W_1^0 = W_2^0$.

372 Si W_1, W_2 son subespacios de $\mathbf{E}$, probar que $(W_1 + W_2)^0 = W_1^0 + W_2^0$.

373 Si W_1, W_2 son subespacios de $\mathbf{E}$ y $W_1 \oplus W_2 = \mathbf{E}$, entonces $\mathbf{E}^* = W_1^0 \oplus W_2^0$.

374 Si $T : \mathbf{E} \to \mathbf{E}$ es un operador lineal y W es un subespacio de $\mathbf{E}$, demostrar que W es T-invariante (cfr. el ejercicio propuesto 308 de este capítulo) si y sólo si W^0 es T^t-invariante, donde T^t es la transformación transpuesta de T definida en el ejercicio propuesto 367 de este apartado.

375 Si $\mathbf{E} = \mathbb{R}^4$ y los funcionales lineales f_i definidos por

$$f_1(x_1,x_2,x_3,x_4) = x_1 + x_2 - x_3 + x_4,$$
$$f_2(x_1,x_2,x_3,x_4) = -x_2 + x_3,$$
$$f_3(x_1,x_2,x_3,x_4) = -x_1 + x_2 - x_3.$$

Encontrar $S \subset \mathbb{R}^4$ tal que $S^0 = \text{gn}(f_1, f_2, f_3)$. Hallar una base y la dimensión de S^0.

376 Sea $V = \text{gn}((1,1,-2,3),(1,2,-1,2),(-1,-2,1,-2))$, hallar una base de W^0.

En los ejercicios 377 a 382 T^t denota la transformación transpuesta de T definida en el ejercicio propuesto 367 de esta sección.

377 Sea el funcional lineal $\phi : \mathbb{R}^2 \to \mathbb{R}$ definido por $\phi(x,y) = 2x - y$. Determinar $T^t(\phi)$ para cada transformación $T : \mathbb{R}^3 \to \mathbb{R}^2$ si:

(a) $T(x,y,z) = (x - y + z, y - 2z)$.

(b) $T(x,y,z) = (2x - y, x + y - z)$.

(c) $T(x,y,z) = (-x + z, x - y - 3z)$.

378 Sean $\mathbf{E} = \mathbf{P}_1$, $\mathbf{F} = \mathbb{R}^2$, con sendas bases $\mathscr{B}_1 = \{1, x\}$, $\mathscr{B}_2 = \{(1,0),(0,1)\}$ y $T : \mathbf{E} \to \mathbf{F}$ definido por

$$T(p) = (p(0) - p(1), 2p(0) + p'(0)).$$

(a) Mostrar que $T \in \mathscr{L}(\mathbf{E}, \mathbf{F})$.

(b) Si $f:\mathbf{F}\to\mathbb{R}$ es el funcional lineal

$$f(x,y)=x-y,$$

encontrar $T^t(f)$

(c) Hallar $[T^t]_{\mathscr{B}_1^*}^{\mathscr{B}_2^*}$ directamente y sin recurrir al resultado del ejercicio propuesto 367.

(d) Encontrar $[T]_{\mathscr{B}_1}^{\mathscr{B}_2}$, transponer y comparar con el resultado del inciso anterior.

379 Si $T_1\in\mathscr{L}(\mathbf{E},\mathbf{F})$ y $T_2\in\mathscr{L}(\mathbf{F},\mathbf{G})$, mostrar que $(T_2\circ T_1)^t=T_1^t\circ T_2^t$.

380 Si $\mathbf{E}$, $\mathbf{F}$ son espacios vectoriales y $T\in\mathscr{L}(\mathbf{E},\mathbf{F})$, mostrar que

$$\operatorname{Ker}(T^t)=(T^t(\mathbf{F}^*))^0.$$

381 Si $\mathbf{E}$, $\mathbf{F}$ son espacios vectoriales, $T\in\mathscr{L}(\mathbf{E},\mathbf{F})$ y $\mathbf{F}$ tiene dimensión finita, utilizar el ejercicio precedente y el ejercicio resuelto 40*(b)* para mostrar que

$$\operatorname{Rang}(T)=\operatorname{Rang}(T^t).$$

382 Si $\mathbf{E}$, $\mathbf{F}$ son espacios vectoriales de dimensiones finitas y $T\in\mathscr{L}(\mathbf{E},\mathbf{F})$ probar que

$$(\operatorname{Ker}(T))^0=T^t(\mathbf{F}^*).$$

383 Sea c_{00} el espacio de sucesiones finitas (cfr. el ejercicio propuesto 357 del cap. 3, pág. 228); es decir,

$$c_{00}=\{(a_n)\in\mathbb{R}^\infty\,|\,a_n=0 \text{ salvo un número finito de índices } n\}.$$

Demostrar que $c_{00}\cong\mathbf{P}$ (el espacio de polinomios).

384 Si $\mathbf{E}$ y $\mathbf{F}$ son espacios vectoriales de dimensiones 2 y 3, respectivamente, y $T_i\in\mathscr{L}(\mathbf{E},\mathbf{F})$, $i=1,\ldots,7$; probar que existen constantes $\alpha_1,\ldots,\alpha_7$, no todas nulas, tales $\sum_{i=1}^7\alpha_iT_i(\vec{u})=\vec{0}_{\mathbf{F}}$ para todo $\vec{u}\in\mathbf{E}$.

385 Sean $\alpha,\beta\in\mathbb{R}$ con $\beta\neq0$ y S el subespacio de $\mathbb{R}^\infty$ de sucesiones (x_n) que satisfacen

$$x_{n+2}+\alpha x_{n+1}+\beta x_n=0\ \forall n$$

Sea $T:S\to\mathbb{R}^2$ definido por

$$T((x_n))=(x_1,x_2).$$

(a) Mostrar que $S<\mathbb{R}^\infty$.

(b) Probar que T es un isomorfismo y determinar la dimensión de S.

(c) Si $\lambda,\mu\in\mathbb{R}$ son dos raíces distintas de la ecuación

$$X^2+\alpha X+\beta=0,$$

mostrar que las sucesiones $u=(\lambda^n)$ y $v=(\beta^n)$ son una base de S.

386 Sean $\mathbf{E}$, $\mathbf{F}$ dos espacios vectoriales y $\mathbf{E}\times\mathbf{F}$ el espacio producto de éstos definido en el ejercicio propuesto 133 del capítulo 3. Mostrar que $\mathbf{E}\times\mathbf{F}\cong\mathbf{F}\times\mathbf{E}$.

387 Si $\mathbf{E}$ es un espacio vectorial y S_1, S_2 son subespacios de $\mathbf{E}$ tales que $\mathbf{E} = S_1 \oplus S_2$, demostrar que $\mathbf{E}\cong S_1\times S_2$.

388 Probar que $\mathbb{R}^m\times\mathbb{R}^n\cong\mathbb{R}^{mn}$.

389 Mostrar que $\mathfrak{M}_{m\times n}\cong\mathbb{R}^{mn}$, sin utilizar el hecho de que todo par de espacios vectoriales que tienen la misma dimensión son isomorfos.

390 Probar que el espacio de matrices simétricas de orden 2 es isomorfo a $\mathbb{R}^3$.

391 Encontrar un isomorfismo $T:\mathbf{P}_n\to\mathbb{R}^{n+1}$.

392 Sea $P\in\mathfrak{M}_n$ una matriz invertible. Mostrar que $T:\mathfrak{M}_n\to\mathfrak{M}_n$ definida por $T(A) = P^{-1}AP$ es un isomorfismo.

393 Sean $\mathbf{E}$, $\mathbf{F}$ espacios vectoriales y $U:\mathbf{E}\to\mathbf{F}$ un isomorfismo. Se define $\Phi:\mathscr{L}(\mathbf{E},\mathbf{E})\to\mathscr{L}(\mathbf{F},\mathbf{F})$ por

$$\Phi(T) = U^{-1}TU$$

mostrar que Φ es un isomorfismo.

394 Sean $\mathbf{E}$, $\mathbf{F}$ espacios vectoriales. Si $T\in\mathscr{L}(\mathbf{E},\mathbf{F})$, $\mathbf{E}$ tiene dimensión finita y $\mathscr{B}$ es una base de $\mathbf{E}$. Probar que T es un isomorfismo si y sólo si $T(\mathscr{B})$ es una base de $\mathbf{F}$.

395 Sean $\mathbf{E}$ un espacio vectorial y U, W dos subespacios de $\mathbf{E}$. Demostrar que

$$(U+W)/W\cong U/(U\cap W).$$

Donde $U+V$ es la suma de subespacios (cfr. el ejercicio resuelto 26 del capítulo 3) y G/H es el espacio cociente (cfr. los ejercicios resueltos 29 y 41 de los capítulos 3 y 5, respectivamente).

396 Sea T una tranformación lineal del espacio $\mathbf{E}$ sobre el espacio $\mathbf{F}$. Si $W<\mathbf{F}$ y $T^{-1}(W)$ es el subespacio imagen inversa de W bajo T (cfr. el ejercicio resuelto 8 de este capítulo), probar que:

(a) $\operatorname{Ker}(T)\subset T^{-1}(W)$.

(b) $T^{-1}(W)/\operatorname{Ker}(T)\cong W$.

397 Sean U, W dos subespacios de un espacio vectorial $\mathbf{E}$ y $K = \{(\vec{u},\vec{v})\in\mathbf{E}\times\mathbf{E}\,|\,\vec{v} = -\vec{u}\}$, demostrar que

$$(U\times W)/K\cong U+W$$

Valores y vectores propios, diagonalización (respuestas en páginas 1090-1092)

398 Sean $\mathbf{E}$ un espacio vectorial, $T:\mathbf{E}\to\mathbf{E}$ un operador lineal, λ un valor propio de T y a un escalar; demostrar que $a\lambda$ es valor propio del operador lineal aT.

399 Sean T un operador lineal en un espacio vectorial $\mathbf{E}$ y λ un valor propio de T con $\vec{u}$ un vector propio correspondiente.

(a) Mostrar que λ^2 es un valor propio de T^2 y que $\vec{u}$ es un vector propio correspondiente a λ^2.

(b) Si $n \in \mathbb{N}$ mostrar que λ^n es un valor propio de T^n y que $\vec{u}$ es un vector propio correspondiente a λ^n.

400 Sea $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ un polinomio. Si $T : \mathbf{E} \to \mathbf{E}$ es un operador lineal en el espacio $\mathbf{E}$, se define el operador lineal

$$p(T) = a_0I + a_1T + a_2T^2 + \cdots + a_nT^n$$

donde, como en ejercicios anteriores, $T^k = \underbrace{T \circ \cdots \circ T}_{k}$ es la composición del operador T consigo mismo k-veces. Si λ es un valor propio de T con $\vec{u}$ un vector propio correspondiente, probar que $\mu = p(\lambda)$ es un valor propio del operador $p(T)$ con $\vec{u}$ un vector propio correspondiente.

401 Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ el operador lineal que rota cada vector $\pi/2$ radianes en sentido contrario a las manecillas del reloj conservando la norma (cfr. el ejercicio resuelto 3 de este capítulo).

(a) Probar que T no tiene valores propios.

(b) Mostrar que todo vector no nulo es vector propio de T^2.

402 Si $T : \mathbf{E} \to \mathbf{E}$ es un operador lineal en el espacio $\mathbf{E}$ y λ^2 es un valor propio del operador T^2, mostrar que λ o $-\lambda$ es un valor propio de T.

403 Sean $C^\infty(0,1)$ el espacio vectorial de funciones $f : (0,1) \to \mathbb{R}$ que tienen derivada de todo orden en cualquier punto del intervalo $(0,1)$ y $T : C^\infty(0,1) \to C^\infty(0,1)$ definido, para cada $f \in C^\infty(0,1)$, por $T(f) = g$ donde $g(x) = xf'(x)$ para todo $x \in (0,1)$. Encontrar los valores y vectores propios correspondientes.

404 Sea el operador lineal $T : \mathfrak{M}_2 \to \mathfrak{M}_2$ definido por $T(A) = A^t$. Encontrar los valores y vectores propios correspondientes.

405 Sea $\mathbf{E} = C(-\infty,\infty)$ el espacio de funciones continuas en todo punto y S el conjunto de funciones $f \in \mathbf{E}$ tales que $\int_{-\infty}^{x} tf(t)dt$ converge para todo $x \in \mathbb{R}$.

(a) Mostrar que S es un subespacio de $\mathbf{E}$.

(b) Sea, para cada $f \in S$, $T(f) = g$ donde $g(x) = \int_{-\infty}^{x} tf(t)dt$. Mostrar que $g \in \mathbf{E}$ y que $T : S \to \mathbf{E}$ así definido es un operador lineal (cfr. el ejercicio resuelto 43 de este capítulo).

(c) Encontrar los valores y vectores propios correspondientes de T.

406 Sea $T : \mathbf{P}_n \to \mathbf{P}_n$ el operador lineal definido por $T(p) = q$ donde $q(x) = p(x+1)$. Encontrar los valores y vectores propios correspondientes de T.

407 Sean $\mathbf{E} = C[0,\pi]$ el espacio de funciones continuas en todo punto del intervalo $[0,\pi]$, S el subconjunto de $\mathbf{E}$ de funciones f dos veces derivables con continuidad que satisfacen $f(0) = 0 = f(\pi)$.

(a) Mostrar que S es un subespacio de $\mathbf{E}$.

(b) Sea $T : S \to \mathbf{E}$ definido por $T(f) = f''$. Probar que T es lineal.

(c) Encontrar los valores y vectores propios correspondientes del operador T.

408 Encontrar los valores y vectores propios correspondientes del operador T definido en el ejercicio propuesto 124 de este capítulo.

409 Sea $T : \mathfrak{M}_2 \to \mathfrak{M}_2$ el operador definido, para cada $A \in \mathfrak{M}_2$, por (cfr. el ejercicio propuesto 103 de este capítulo)

$$T(A) = \frac{A + A^t}{2}.$$

Hallar los valores propios y vectores propios correspondientes de T.

410 Sea $T : \mathfrak{M}_2 \to \mathfrak{M}_2$ el operador definido, para cada $A \in \mathfrak{M}_2$, por (cfr. el ejercicio propuesto 107 de este capítulo)

$$T(A) = \frac{A - A^t}{2}.$$

Hallar los valores y vectores propios correspondientes de T.

411 Sean T un operador lineal en un espacio $\mathbf{E}$, λ y μ ($\lambda \neq \mu$) valores propios de T con sendos vectores propios $\vec{u}$ y $\vec{v}$. Mostrar que si $a\vec{u} + b\vec{v}$ es un vector propio de T, entonces $a = 0$ o $b = 0$.

412 Probar que si T es un operador lineal en un espacio $\mathbf{E}$ tal que todo vector no nulo de $\mathbf{E}$ es un vector propio de T, entonces éste tiene la forma $T(\vec{u}) = c\vec{u}$.

413 Sean $T, R : \mathbf{E} \to \mathbf{E}$ dos operadores lineales en el espacio $\mathbf{E}$. Se supone que $\vec{u}$ es un vector propio de T correspondiente a un valor propio λ de este operador y que $R(\vec{u}) \neq \vec{0}_{\mathbf{E}}$, ¿$R(\vec{u})$ es vector propio del operador T correspondiente al mismo valor propio λ?

414 Sea $T : \mathbf{E} \to \mathbf{E}$ un operador lineal en el espacio vectorial $\mathbf{E}$.

(a) Demostrar que T es inyectivo si y sólo si $\lambda = 0$ no es valor propio de T.

(b) Si $\mathbf{E}$ tiene dimensión finita, mostrar que T es invertible si y sólo si $\lambda = 0$ no es valor propio de T.

(c) Si T es invertible probar que λ es valor propio de T si y sólo si λ^{-1} es valor propio de T^{-1}.

415 Probar que los valores propios de una matriz cuadrada triangular superior o triangular inferior son los elementos de la diagonal.

416 Si A es una matriz cuadrada demostrar que A y A^t tienen el mismo polinomio característico y por tanto los mismos valores propios.

417 Sean $\mathbf{E}$ un espacio vectorial de dimensión finita, $T : \mathbf{E} \to \mathbf{E}$ un operador lineal y $T^t : \mathbf{E}^* \to \mathbf{E}^*$ el operador transpuesto de T definido en el ejercicio propuesto 367 de esta sección. Probar que T y T^t tienen los mismos valores propios.

418 Sean $\mathbf{E}$ es un espacio vectorial de dimensión finita, $T : \mathbf{E} \to \mathbf{E}$ un operador lineal y $T^* : \mathbf{E} \to \mathbf{E}$ el operador adjunto de T definido en el ejercicio resuelto 38 de este capítulo; mostrar que T y T^* tienen los mismos valores propios (cfr. el ejercicio resuelto 39 de este capítulo).

419 Probar que si $A, B \in \mathfrak{M}_n$ y una de ellas es invertible, entonces AB y BA son similares y por tanto tienen los mismos valores propios.

420 Sean A, B dos matrices cuadradas de orden n.

(a) Probar que si $AB - I_n$ es invertible entonces $BA - I_n$ es también invertible y que

$$(BA - I_n)^{-1} = B(AB - I)^{-1}A - I_n.$$

(b) Mostrar que AB y BA tienen los mismos valores propios.

421 Si $A = \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, calcular los valores propios de AB y BA. Por el ejercicio precedente estos productos deben tener los mismos valores propios, ¿tienen los mismos vectores propios correspondientes?

En los ejercicios 422 a 440, para la matriz A: *(i)* encontrar los valores propios; *(ii)* hallar los vectores y espacios propios correspondientes; *(iii)* calcular las dimensiones de los espacios propios.

422 $A = \begin{bmatrix} 3 & -2 \\ 4 & -3 \end{bmatrix}$.

423 $A = \begin{bmatrix} -1 & 0 \\ -3 & -1 \end{bmatrix}$.

424 $A = \begin{bmatrix} 5 & -4 \\ 8 & -7 \end{bmatrix}$.

425 $A = \begin{bmatrix} -5 & 12 \\ -2 & 5 \end{bmatrix}$.

426 $A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$.

427 $A = \begin{bmatrix} 0 & -1 \\ 2 & 3 \end{bmatrix}$.

428 $A = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$.

429 $A = \begin{bmatrix} -1 & 1 \\ 2 & 4 \end{bmatrix}$.

430 $A = \begin{bmatrix} 7 & 16 \\ -3 & -7 \end{bmatrix}$.

431 $A = \begin{bmatrix} -3 & 2 & 0 \\ -4 & 3 & 0 \\ 4 & -2 & 1 \end{bmatrix}$.

432 $A = \begin{bmatrix} -3 & 2 & 0 \\ -4 & 3 & 0 \\ 3 & -3 & -1 \end{bmatrix}$.

433 $A = \begin{bmatrix} 1 & -3 & 4 \\ 0 & 4 & -2 \\ 0 & 1 & 1 \end{bmatrix}$.

434 $A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -6 & -4 \\ -1 & 8 & 6 \end{bmatrix}$.

435 $A = \begin{bmatrix} 5 & 0 & -3 \\ 6 & -1 & -3 \\ 6 & 0 & -4 \end{bmatrix}$.

436 $A = \begin{bmatrix} -1 & -2 & 1 \\ -2 & -1 & 1 \\ 2 & -4 & 0 \end{bmatrix}$.

437 $A = \begin{bmatrix} 5 & 0 & 6 \\ 2 & 1 & 4 \\ 0 & -1 & 2 \end{bmatrix}$.

438 $A = \begin{bmatrix} 1 & 0 & 0 \\ -3 & -2 & 2 \\ 2 & 1 & 2 \end{bmatrix}$.

439 $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & -3 & 1 \\ 0 & 1 & -1 \end{bmatrix}$.

440 $A = \begin{bmatrix} 6 & -1 & -2 \\ 4 & 1 & -2 \\ 6 & -2 & -1 \end{bmatrix}$.

En los ejercicios 441 a 463 determinar si la matriz A o el operador T es o no diagonalizable. En caso afirmativo encontrar una diagonalización para la matriz o el operador en cuestión.

441 A es la matriz del ejercicio propuesto 422 de este capítulo.

442 A es la matriz del ejercicio propuesto 424 de este capítulo.

443 $A = \begin{bmatrix} 2 & -2 \\ 2 & 6 \end{bmatrix}$.

444 A es la matriz del ejercicio propuesto 425 de este capítulo.

445 A es la matriz del ejercicio propuesto 426 de este capítulo.

446 $A = \begin{bmatrix} 5 & 3 & 0 \\ -3 & -1 & 0 \\ 1 & 1 & 3 \end{bmatrix}$.

447 A es la matriz del ejercicio propuesto 431 de este capítulo.

448 $A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ -1 & 0 & 4 \end{bmatrix}$.

449 $A = \begin{bmatrix} 8 & 3 & -3 \\ -6 & -1 & 6 \\ 0 & 0 & 5 \end{bmatrix}$.

450 A es la matriz del ejercicio propuesto 432 de este capítulo.

451 $T(x,y,z) = (-y+2z, 3x+4y-2z, -x+4z)$.

452 $T(x,y,z) = (-7x+15y-11z, -6x+12y-7z, z)$.

453 $T : \mathfrak{M}_2 \to \mathfrak{M}_2$, $T(A) = A^t$.

454 $T : \mathfrak{M}_2 \to \mathfrak{M}_2$, $T(A) = \frac{A+A^t}{2}$.

455 $T : \mathfrak{M}_2 \to \mathfrak{M}_2$, $T(A) = \frac{A-A^t}{2}$.

456 A es la matriz del ejercicio propuesto 438 de este capítulo.

457 A es la matriz del ejercicio propuesto 439 de este capítulo.

458 $T(x,y,z) = (2x-7y, 2x-6y, -x-y+z)$.

459 $A = \begin{bmatrix} 2 & -6 & 0 \\ 2 & -5 & 0 \\ -3 & 5 & -2 \end{bmatrix}$.

460 $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ -4 & 4 & -2 & -3 \\ 3 & -3 & 3 & 4 \end{bmatrix}$.

461 $A = \begin{bmatrix} -4 & 0 & -2 & -3 \\ 0 & -2 & 0 & 0 \\ 7 & 0 & 5 & 3 \\ 2 & 0 & 2 & 1 \end{bmatrix}$.

462 $A = \begin{bmatrix} 5 & 0 & 4 & 0 \\ 0 & -3 & 0 & 0 \\ -8 & 0 & -7 & 0 \\ 4 & 0 & 4 & 1 \end{bmatrix}$.

463 $A = \begin{bmatrix} 2 & 0 & 2 & -2 \\ -4 & 2 & -3 & 2 \\ -1 & 0 & -1 & 2 \\ -1 & 0 & 0 & 1 \end{bmatrix}$.

464 Sea $T : \mathbf{P}_n \to \mathbf{P}_n$ definido por $T(p) = q$ donde $q(x) = p(x) + xp'(x)$.

(a) Probar que T es un operador lineal.

(b) Encontrar los valores propios de T.

(c) Determinar si T es diagonalizable y en caso positivo hallar una diagonalización para T.

En los ejercicios 465 a 467 para la transformación $T : \mathbf{E} \to \mathbf{E}$: *(i)* probar que T es lineal; *(ii)* determinar si T es diagonalizable y en caso afirmativo encontrar una base $\mathscr{B}$ de $\mathbf{E}$ tal que $[T]_{\mathscr{B}}$ sea diagonal.

465 $T : \mathbf{P}_3 \to \mathbf{P}_3$, $T(p) = p' + p''$.

466 $T : \mathbf{P}_2 \to P_2$, $T(a_0 + a_1x + a_2x^2) = a_2 + a_1x + a_0x^2$.

467 $T : \mathbf{P}_2 \to \mathbf{P}_2$, $T(a_0 + a_1x + a_2x^2) = a_0 + \frac{1}{2}a_1x + \frac{1}{3}a_2x^2$.

468 Una matriz $A \in \mathfrak{M}_n$ es escalar si $A = \lambda I_n$ para algún escalar λ. Demostrar que si B es una matriz de orden n y tiene un único valor propio, entonces B es una matriz escalar.

469 ¿Es diagonalizable la matriz $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{bmatrix}$?

470 Encontrar condiciones necesarias y suficientes, respecto a los parámetros a, b, c, d, para que la matriz $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ sea diagonalizable (sobre $\mathbb{R}$).

471 Determinar las condiciones sobre los parámetros $a, b \in \mathbb{R}$ para que la matriz

$$\begin{bmatrix} 2a-b & 0 & 2a-2b \\ 1 & a & 2 \\ -a+b & 0 & -a+2b \end{bmatrix}$$

sea diagonalizable.

472 Sea

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ a_1 & b & 2 & 0 \\ a_2 & b_1 & c & 2 \end{bmatrix}.$$

Encontrar condiciones sobre los parámetros $a_1, a_2, b, b_1 \in \mathbb{R}$ para que la matriz A sea diagonalizable.

En los ejercicios 473 a 484 se consideran valores propios, vectores propios y diagonalización sobre el campo $\mathbb{C}$. Para la matriz A: *(i)* encontrar los valores propios; *(ii)* bases para los espacios propios; *(iii)* determinar si la matriz A es diagonalizable y en caso afirmativo hallar una diagonalización para la matriz A.

473 $A = \begin{bmatrix} i & 5 \\ 1 & -i \end{bmatrix}$.

474 $A=\begin{bmatrix} 1-3i & 2 \\ 4 & 1+3i \end{bmatrix}$.

475 $A=\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$.

476 $A=\begin{bmatrix} 2 & -2 \\ 4 & -2 \end{bmatrix}$.

477 $A=\begin{bmatrix} 2+i & 3 \\ 0 & 2+i \end{bmatrix}$.

478 $A=\begin{bmatrix} 3i & -i \\ 4i & -i \end{bmatrix}$.

479 $A=\begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

480 $A=\begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & -i \end{bmatrix}$.

481 $A=\begin{bmatrix} i & -1+i & -2+i \\ 0 & -3+2i & -3+i \\ 0 & 6-2i & 6-i \end{bmatrix}$.

482 $A=\begin{bmatrix} -i & i & 0 \\ -2-2i & 1+2i & 0 \\ -3-3i & 3i & 1 \end{bmatrix}$.

483 $A=\begin{bmatrix} -2i & -1-i & 0 \\ 2+2i & 3+i & 0 \\ -1-i & -1-i & 1-i \end{bmatrix}$.

484 $A=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.

En los ejercicios 485 a 489 hallar A^n para todo n, $\det(A)$ y $\operatorname{tra}(A)$ para la matriz diagonalizable A utilizando las propiedades de las matrices diagonalizables (teorema 5.34, pág. 481) y de los valores propios de una matriz (teorema 5.36, pág. 486).

485 $A=\begin{bmatrix} 8 & 3 & -3 \\ -6 & -1 & 6 \\ 0 & 0 & 5 \end{bmatrix}$.

486 $A=\begin{bmatrix} 0 & -3 & -2 \\ 2 & 5 & 2 \\ -2 & -6 & -3 \end{bmatrix}$.

487 $A = \begin{bmatrix} -3i & 0 & 2i \\ 4i & i & -2i \\ -4i & 0 & 3i \end{bmatrix}$.

488 $A = \begin{bmatrix} 1 & 0 & -1-3i & 2i \\ 0 & 1 & -3-i & 2 \\ 0 & 0 & -3i & 2i \\ 0 & 0 & -4i & 3i \end{bmatrix}$.

489 $A = \begin{bmatrix} 1 & 0 & 2 & 2 \\ 2 & 1 & -4 & -2 \\ 2 & 0 & 2 & 3 \\ 0 & 0 & -3 & -2 \end{bmatrix}$.

490 Sea **E** un espacio vectorial complejo con producto interior $\langle\cdot,\cdot\rangle$ y $\vec{u}_0$ un vector de **E**. Demostrar que la función $f:\mathbf{E}\to\mathbb{C}$ definida por

$$f(\vec{v}) = \langle\vec{v},\vec{u}_0\rangle$$

es lineal.

491 Sea **E** un espacio vectorial complejo de dimensión finita y $f:\mathbf{E}\to\mathbb{C}$ un funcional lineal. Mostrar que existe un $\vec{u}_f\in\mathbf{E}$ tal que

$$f(\vec{u}) = \langle\vec{u},\vec{u}_f\rangle$$

para todo $\vec{u}\in\mathbf{E}$.

492 Sea $T:\mathbb{C}^3\to\mathbb{C}$ el funcional lineal definido por

$$f(z_1,z_2,z_3) = (2+i)z_1 - z_2 + (2-i)z_3.$$

Hallar $\vec{u}_0\in\mathbb{C}^3$ tal que

$$f(\vec{z}) = \langle\vec{z},\vec{u}_0\rangle$$

para todo $\vec{z}\in\mathbb{C}^3$, donde $\langle(z_1,z_2,z_3),(w_1,w_2,w_3)\rangle = \vec{z}\cdot\vec{w} = \sum_{i=1}^{3} z_i\overline{w_i}$ es el producto canónico en $\mathbb{C}^3$ (cfr. el ejemplo 5.60, pág. 492).

493 Sean **E** un espacio vectorial de dimensión finita sobre los números complejos, $\langle\cdot,\cdot\rangle$ un producto interior en **E** y $T:\mathbf{E}\to\mathbf{E}$ un operador lineal. Probar que existe un operador lineal $T^*:\mathbf{E}\to\mathbf{E}$ tal que

$$\langle T(\vec{u}),\vec{v}\rangle = \langle\vec{u},T^*(\vec{v})\rangle \quad \forall\vec{u},\vec{v}\in\mathbf{E}.$$

Al operador T^* se le dice el operador adjunto de T.

494 Sean **E** y T como en el ejercicio anterior. Probar que si $A = [T]_{\mathscr{B}}$ es la representación matricial de T relativa a una base ortonormal $\mathscr{B}$ y $B = [T^*]_{\mathscr{B}}$ es la representación matricial del operador adjunto T^*, entonces $B = \overline{A^t}$, donde $\overline{M}$ significa la matriz que tiene por componentes los conjugados de las componentes de la matriz M.

495 Sea $T : \mathbb{C}^3 \to \mathbb{C}^3$ el operador lineal definido por

$$T(z_1, z_2, z_3) = (iz_1 + (2-3i)z_2, 2z_1 + 3iz_2, z_2 - iz_3).$$

Hallar T^* si se considera el producto interior $\langle \vec{z}, \vec{w} \rangle = \vec{z} \cdot \overline{\vec{w}}$ (cfr. el ejemplo 5.60).

496 Sea T un operador lineal en un espacio $\mathbf{E}$. Probar que el subespacio S es T-invariante ($T(S) \subset S$) si:

(a) $S = \{\vec{0}_{\mathbf{E}}\}$.

(b) $S = \mathbf{E}$.

(c) $S = T(\mathbf{E})$.

(d) $S = \operatorname{Ker}(T)$.

(e) $S = E_\lambda$ para cualquier valor propio λ del operador T.

En los ejercicios 497 a 500 determinar si el subespacio S es T-invariante para la transformación $T : \mathbf{E} \to \mathbf{E}$; esto es, si $T(S) \subset S$.

497 $T : \mathbf{P}_3 \to \mathbf{P}_3$, $T(p) = p'$, $S = \mathbf{P}_2$.

498 $T : \mathbf{P} \to \mathbf{P}$, $T(p) = q$ donde $q(x) = xp(x)$, $S = \mathbf{P}_2$.

499 $T : C[0,1] \to C[0,1]$, $T(f) = g$ donde $g(x) = x\int_0^1 f(t)dt$, $S = \{g \in C[0,1] \,|\, g(x) = a_0 + a_1 x,\ a_0, a_1 \in \mathbb{R}\}$.

500 $T : \mathfrak{M}_2 \to \mathfrak{M}_2$, $T(A) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} A$, S es subespacio de matrices simétricas.

501 Probar que si $T : \mathbf{E} \to \mathbf{E}$ es un operador lineal en el espacio $\mathbf{E}$ y S_α, $\alpha \in \Lambda$, es una colección de subespacios de $\mathbf{E}$, entonces $S = \bigcap_{\alpha \in \Lambda} S_\alpha$ es un subespacio T-invariante; es decir, $T(S) \subset S$ (cfr. el ejercicio resuelto 30 del capítulo 3, pág. 190).

En los ejercicios 502 a 504 encontrar una base para el subespacio T-cíclico, definido en el ejercicio resuelto 48 de este capítulo, generado por el vector $\vec{u}$.

502 $T : \mathbf{P}_3 \to \mathbf{P}_3$, $T(p) = p''$, $\vec{u} = x^3$.

503 $T : \mathfrak{M}_3 \to \mathfrak{M}_3$, $T(A) = A^t$, $\vec{u} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

504 $T : \mathbf{P}_2 \to \mathbf{P}_2$, $T(p) = p'$,$\vec{u} = x^2$.

En los ejercicios 505 a 514 encontrar el polinomio característico y el polinomio mínimo de la matriz A. En algunos casos se debe considerar trabajar sobre el campo $\mathbb{C}$.

505 $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

506 $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

507 $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

508 $A = \begin{bmatrix} 4 & 3 & 0 \\ -6 & -5 & 0 \\ 3 & 3 & 1 \end{bmatrix}$.

509 $A = \begin{bmatrix} -4 & 0 & -3 \\ 3 & -1 & 3 \\ 6 & 0 & 5 \end{bmatrix}$.

510 $A = \begin{bmatrix} -5 & 0 & 4 \\ 4 & -1 & -4 \\ -8 & 0 & 7 \end{bmatrix}$.

511 $A = \begin{bmatrix} -1 & 0 & 1 & -1 \\ 2 & 1 & -1 & 1 \\ -3 & 0 & 3 & -1 \\ 3 & 0 & -1 & 3 \end{bmatrix}$.

512 $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 0 & 2 \\ -3 & -2 & -1 \end{bmatrix}$.

513 $A = \begin{bmatrix} 0 & -1 & 2 & -1 \\ 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$.

514 $A = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 3 & 1 & 1 & 1 \\ 2 & -5 & -2 & -1 \end{bmatrix}$.

515 Demostrar el teorema de Cayley-Hamilton (cfr. el ejercicio resuelto 50 de este capítulo) para una matriz 2×2, $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, utilizando directamente la fórmula

$$p_A(\lambda) = \lambda^2 - \text{tra}(A)\lambda + \det(A)$$

para el polinomio característico de A para matrices de orden 2.

Aplicaciones, uso de tecnología, métodos numéricos

6 Aplicaciones

Se han seleccionado unas cuantas de las aplicaciones que tiene la materia de álgebra lineal para formar parte de este capítulo. Van desde lo elemental hasta aplicaciones más sofisticadas. Todas deben considerarse como introducciones a temas que son bastos y complejos; sin embargo, algunas de ellas se han tratado con suficiente detalle para que este libro resulte, en la medida de lo posible, un texto autocontenido. Las primeras aplicaciones ilustran conceptos muy simples y cuestiones bastante elementales de álgebra lineal; pero las demás involucran conceptos que en sí requieren de una formulación conceptual más complicada. Las secciones 6.1, 6.2, 6.4, 6.5 y 6.6 necesitan, esencialmente, el conocimiento de matrices y sistemas lineales. La sección 6.3 requiere, además, saber los conceptos básicos de matrices invertibles, así que pueden abordarse inmediatamente que se hayan cubierto esos temas. Para lograr una mayor flexibilidad en el manejo de este capítulo se hacen patentes, al inicio de cada sección y donde es necesario, los requisitos de álgebra lineal –mencionando los apartados donde están incluidos en el libro– y los conocimientos de otras ramas de las matemáticas que son necesarios para su comprensión. Por tanto, su estudio puede llevarse a cabo en el orden que convenga y, no necesariamente, en el orden en el que aparecen los tópicos. La última sección contiene una serie de ejercicios propuestos al lector para su resolución.

6.1 Matrices de incidencia y teoría de grafos

Requisitos: *Operaciones con matrices: 1.1.1 a 1.1.3*

Una matriz de **incidencia** es una matriz cuadrada cuyas entradas son ceros o unos y, por conveniencia, todas las componentes de la diagonal son nulas.

► **Ejemplo 6.1**

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

es una matriz de incidencia.◄

Estas matrices tienen una interpretación muy sencilla que es de gran utilidad en varias aplicaciones. Supongamos que se tienen n personas, $P_1, \ldots, P_n$, con ciertos dispositivos de comunicación, de manera

que algunas personas pueden enviar (transmitir) mensajes a otras. Formamos una matriz $A = [a_{ij}]$ para este esquema definiendo

$$a_{ij} = \begin{cases} 1, \text{ si } i \neq j \text{ y } P_i \text{ puede transmitir a } P_j; \\ 0 \quad \text{en otro caso.} \end{cases}$$

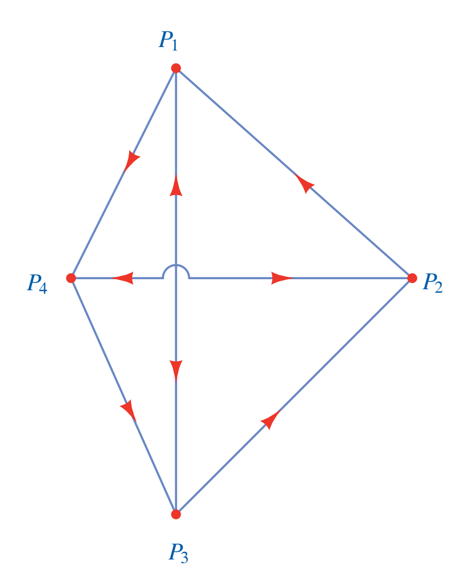


Figura 6-1 •

Entonces, A es una matriz de incidencia. Note que si $a_{ij} = 1$ no necesariamente $a_{ji} = 1$; esto es, si la persona P_i puede enviar mensajes a la persona P_j, no forzosamente la persona P_j puede transmitir a la persona P_i. Una forma de representar las relaciones de comunicación de este tipo es por medio de un diagrama llamado **grafo dirigido** o **digrafo.** Por ejemplo, la figura 6-1 muestra el digrafo de la matriz de incidencia del ejemplo 6.1. En él los individuos P_i, $i = 1, \ldots, 4$, están representados por los puntos P_i que se llaman **vértices** del digrafo y las líneas que unen ciertos vértices se llaman **aristas**; las flechas indican direcciones hacia donde es posible transmitir comunicación.

Aunque en el digrafo de la figura 6-1 P_1 no se puede comunicar directamente con P_2, existe una cadena o trayectoria a través de la cual puede enviar mensajes a P_2, por ejemplo, $P_1 \to P_3 \to P_2$. Se dice que ésta es una 2-cadena porque para transmitir un mensaje es necesario hacerlo en dos etapas, esto es, recorrer una trayectoria que tiene dos aristas con un vértice intermedio. Otra trayectoria por la que P_1 puede transmitir a P_2 es la 3-cadena $P_1 \to P_4 \to P_3 \to P_2$. En general, en un digrafo cualquiera, una k-**cadena**, es una sucesión de la forma $P_i \to P_l \to \cdots \to P_j$ por medio de la cual P_i puede enviar mensajes a P_j, donde esta cadena contiene $k+1$ vértices y k aristas.

Supongamos ahora que un digrafo tiene $n > 1$ vértices y estamos interesados en encontrar el número de formas distintas en las que un individuo P_i puede transmitir al individuo P_j a través de una 2-cadena; es decir, el número total de 2-cadenas con vértice inicial P_i y vértice final P_j o, de manera equivalente, el número de formas distintas en las que P_i puede transmitir a P_j con un intermediario, exactamente, $P_i \to P_l \to P_j$. Representemos por $(A^2)_{ij}$ la componente en la fila i y la columna j de la matriz A^2, donde $A = [A_{ij}]$ es la matriz de incidencia del digrafo. Entonces

$$\begin{aligned}(A^2)_{ij} &= \begin{bmatrix} A_{i1} & \cdots & A_{in} \end{bmatrix} \begin{bmatrix} A_{1j} \\ \vdots \\ A_{nj} \end{bmatrix} \\ &= \sum_{l=1}^{n} A_{il} A_{lj}.\end{aligned}$$

Puesto que A es una matriz de incidencia, los términos de la precedente sumatoria son iguales a 1, o iguales a 0; y ya que $A_{il}A_{lj} = 1$ si y sólo si P_i puede transmitir a P_l y P_l puede transmitir a P_j, se tiene que el resultado de esta sumatoria es el número total de 2-cadenas con vértice inicial P_i y vértice final P_j. La generalización de este resultado es inmediata, se hace patente en el siguiente teorema y su demostración se deja de ejercicio al lector.

Teorema 6.1 *Sea un digrafo con $n > 1$ vértices y matriz de incidencia A. Si $r \geq 1$ es un número entero y $(A^r)_{ij}$ representa la componente ij de la matriz A^r, entonces $(A^r)_{ij}$ es el número de r-cadenas con vértice inicial P_i y vértice final P_j o, equivalentemente, el número de formas distintas en las que P_i puede transmitir a P_j con $r-1$ intermediarios.*

► **Ejemplo 6.2** Consideremos el digrafo que se muestra en la figura 6-1 cuya matriz de incidencia es la matriz del ejemplo 6.1; esto es,

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

Entonces,

$$A^2 = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \end{bmatrix}$$

Así, el individuo P_1 puede transmitir en dos formas distintas al individuo P_2, con un relevo (intermediario) de por medio, las 2-cadenas: $P_1 \to P_4 \to P_2$ y $P_1 \to P_3 \to P_2$; el individuo P_2 puede transmitir al individuo P_3 en dos formas distintas con un relevo de por medio: $P_2 \to P_4 \to P_3$ y $P_2 \to P_1 \to P_3$; sólo hay una 2-cadena con vértice inicial P_1 y vértice final P_4: $P_1 \to P_2 \to P_4$ (es decir, sólo hay una forma de transmisión de P_1 a P_4 en dos etapas); hay cero 2-cadenas con vértice inicial P_4 y vértice final P_3; etc., lo cual se puede constatar observando la figura 6-1.◄

► **Ejemplo 6.3** Si A es la matriz de incidencia del ejemplo anterior,

$$A^3 = \begin{bmatrix} 3 & 1 & 1 & 3 \\ 3 & 3 & 1 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 1 & 3 & 3 \end{bmatrix}.$$

Entonces $(A^3)_{14} = 3$, así que existen tres 3-cadenas con vértice inicial P_1 y vértice final P_4:

$$\begin{array}{ccccccc} P_1 & \to & P_3 & \to & P_2 & \to & P_4, \\ P_1 & \to & P_4 & \to & P_2 & \to & P_4 \quad \text{y} \\ P_1 & \to & P_3 & \to & P_1 & \to & P_4. \end{array}$$ ◄

Es evidente que el teorema precedente incluye en el conteo r-cadenas que no son, en general, de interés en la práctica, son aquellas en las que se repiten vértices; estas cadenas se llaman **redundantes**. En el ejemplo anterior, las 3-cadenas

$$\begin{array}{ccccccc} P_1 & \to & P_4 & \to & P_2 & \to & P_4 \text{ y} \\ P_1 & \to & P_3 & \to & P_1 & \to & P_4 \end{array}$$

son 3-cadenas redundantes.

Si A es una matriz de incidencia, entonces, por el teorema 6.1 $(A^k)_{ij}$, $k = 1, \ldots, r$, es el número de k-cadenas con vértice inicial P_i y vértice final P_j, luego $(A + A^2 + \cdots + A^r)_{ij}$ es el número de cadenas con a lo más r aristas con vértice inicial P_i y vértice final P_j. Hemos probado así el siguiente corolario.

Corolario 6.1 *Si un digrafo tiene n vértices P_i, A es su matriz de incidencia y $r \geq 1$ es un entero, entonces*

$$(A + A^2 + \cdots + A^r)_{ij}$$

es el número de formas distintas que P_i puede transmitir a P_j en a lo más r etapas.

▶ **Ejemplo 6.4** Para la matriz de incidencia

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}:$$

$$A^2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{y}$$

$$A^3 = \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}.$$

Entonces

$$(A + A^2 + A^3)_{12} = 3$$

es el número de formas en las que P_1 puede transmitir a P_2 en a lo más 3 etapas. La figura 6-2 contiene el digrafo de esta matriz, donde se puede observar que las k-cadenas, con $k = 1, 2, 3$, con vértice inicial P_1 y vértice final P_2 son:

$$\begin{array}{lllllll} P_1 & \to & P_2, & & & & \\ P_1 & \to & P_2 & \to & P_1 & \to & P_2, \\ P_1 & \to & P_2 & \to & P_3 & \to & P_2. \end{array}$$ ◀

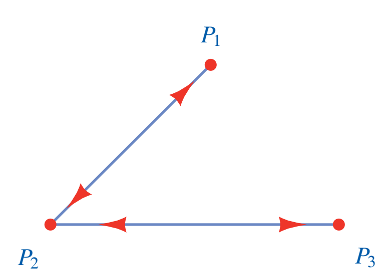


Figura 6-2 • Digrafo de la matriz.

Un problema importante que surge en la teoría de grafos es determinar subconjuntos maximales de vértices que estén intercomunicados. Un subconjunto S de un digrafo es un **clan** si:

1. S contiene 3 o más vértices.
2. Si $P_i, P_j \in S$, entonces $P_i \to P_j$ y $P_j \to P_i$; es decir, cualquier par de vértices en S se pueden comunicar directamente uno con otro.
3. No existe un subconjunto propio del digrafo que satisfaga las dos condiciones precedentes y contenga propiamente a S.

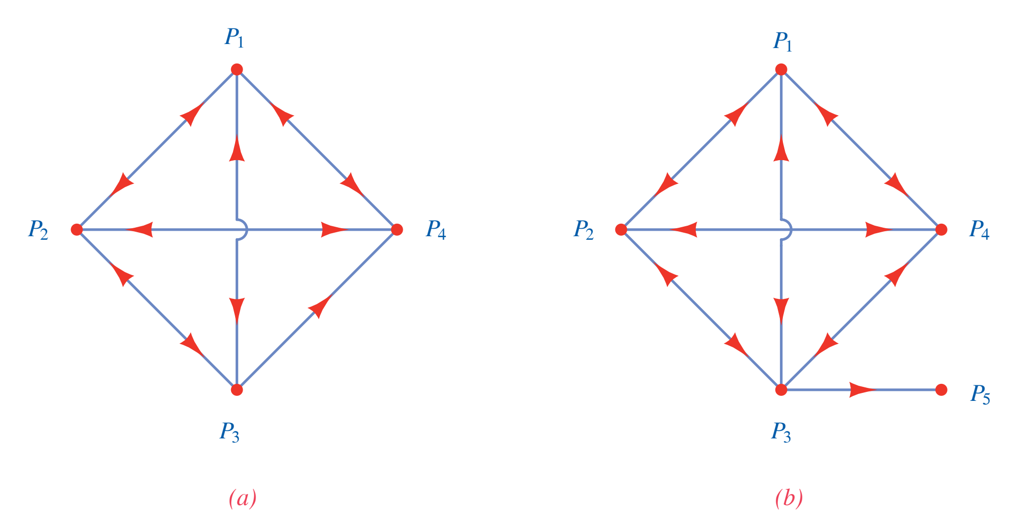


Figura 6-3 •

► **Ejemplo 6.5** En la figura 6-3*(a)* $S_1 = \{P_1, P_2, P_3\}$ y $S_2 = \{P_1, P_2, P_4\}$ son clanes. En la misma figura *(b)* ni $\{P_1, P_2, P_3\}$ ni $\{P_1, P_2, P_4\}$ son clanes, pues ambos subconjuntos están contenidos en $S = \{P_1, P_2, P_3, P_4\}$, el cual sí es un clan (el único) del digrafo.◄

Supongamos ahora que se tiene un digrafo con matriz de incidencia A, se define la matriz $\mathscr{A}(A) = [b_{ij}]$ donde $b_{ij} = 1$ si P_i y P_j pueden transmitir mensajes directos en forma mutua; es decir, en una etapa y en las dos direcciones; y $b_{ij} = 0$ en otro caso. Llamaremos a $\mathscr{A}(A)$ la matriz asociada de la matriz A.

► **Ejemplo 6.6** Hallar la matriz de incidencia del digrafo de la figura 6-3*(a)* y su matriz asociada.◄

Solución De la figura 6-3*(a)* se desprende que

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

y

$$\mathscr{A}(A) = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$ ✔

El siguiente teorema no es difícil de probar y su demostración se deja de ejercicio al lector.

Teorema 6.2 *Si un digrafo tiene* n *vértices con matriz de incidencia A y matriz asociada $B = \mathscr{A}(A)$, entonces P_i pertenece a un clan del digrafo si y sólo si $(B^3)_{ii} > 0$.*

► **Ejemplo 6.7** Consideremos el digrafo contenido en la figura 6-3*(a)*. Por el ejemplo anterior,

$$B = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

ya que

$$B^3 = \begin{bmatrix} 4 & 5 & 5 & 5 \\ 5 & 4 & 5 & 5 \\ 5 & 5 & 2 & 2 \\ 5 & 5 & 2 & 2 \end{bmatrix},$$

todos los vértices P_1, P_2, P_3 y P_4 están en un clan. Efectivamente, en el ejemplo 6.5 vimos que este digrafo tiene dos clanes: $S_1 = \{P_1, P_2, P_3\}$ y $S_2 = \{P_1, P_2, P_4\}$.◄

► **Ejemplo 6.8** Ahora consideremos el digrafo contenido en la figura 6-3*(b)*, su matriz de incidencia es

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

mientras que la matriz asociada a A es

$$B = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Puesto que

$$B^3 = \begin{bmatrix} 6 & 7 & 7 & 7 & 0 \\ 7 & 6 & 7 & 7 & 0 \\ 7 & 7 & 6 & 7 & 0 \\ 7 & 7 & 7 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

se desprende que todos los vértices, excepto P_5, están en un clan. De hecho, por el ejemplo 6.5, sabemos que $\{P_1, P_2, P_3, P_4\}$ es el único clan de este digrafo.◄

Dominancia total

Supongamos ahora que en un digrafo queremos representar una relación entre ciertos individuos P_i que tiene que ver con el dominio, o influencia directa, que algunos ejercen sobre otros. Esta relación puede ser, por ejemplo, de liderazgo o de algún tipo de influencia sicológica o nuevamente de transmisión de información, de tal manera que si un individuo ejerce dominio sobre otro, éste no lo hace sobre el

Figura 6-4 • Digrafo.

primero. Si se supone además que todo individuo domina, o es dominado, en forma directa por los demás miembros del grupo, entonces se dice que la relación es de **dominancia total**. Por tanto, en el digrafo todo par de vértices P_i y P_j están conectados en la forma $P_i \to P_j$ o en la forma $P_j \to P_i$, pero no en ambas. Luego, la matriz de incidencia, $A = [a_{ij}]$, para una relación de dominancia total tienen una característica más: $a_{ij} = 1$ si y sólo si $a_{ji} = 0$. Como antes, la relación de dominancia puede ejercerse a través de k-cadenas, esto es, si $P_i \to P_j$ (y, por tanto, $P_j \nrightarrow P_i$), puede existir una k-cadena tal que $P_j \to \cdots \to P_l \to P_i$. Por ejemplo, en el digrafo contenido en la figura 6-4, $P_4 \to P_2$ y $P_2 \nrightarrow P_4$, pero $P_2 \to P_3 \to P_4$, esto es, P_2 establece dominio sobre P_4 en dos etapas.

► **Ejemplo 6.9** Consideremos el digrafo mostrado en la figura 6-4. Entonces, la matriz de incidencia es

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}.$$

Por tanto, la relación que representa el digrafo de la figura 6-4 es de dominancia total, ya que: las componentes de a_{ij} de A son ceros o unos, $a_{ii} = 0$ para todo i, $a_{ij} = 1 \Leftrightarrow a_{ji} = 0$ para todo $i \neq j$. ◄

Se puede probar que si A es la matriz de incidencia de una relación de dominancia total, entonces $A + A^2$ tiene una fila (columna) con todos sus elementos positivos excepto el de la diagonal y, por el corolario 6.1, se desprende, entonces, que siempre existe un individuo P_i que domina a todos los demás (o es dominado por todos los demás) en a lo sumo dos etapas.

► **Ejemplo 6.10** Si A es la relación de dominancia cuya matriz de incidencia es la matriz A del ejemplo precedente, entonces,

$$A + A^2 = \begin{bmatrix} 0 & 2 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 2 & 1 \\ 1 & 2 & 2 & 0 & 1 \\ 2 & 2 & 2 & 2 & 0 \end{bmatrix}.$$

Luego, los individuos P_1, P_3, P_4 y P_5 ejercen su influencia en todos los demás en, a lo sumo, dos etapas; mientras que P_1, P_2, P_3 y P_4 son dominados por todos los demás en a lo más dos etapas. El lector puede verificar estas afirmaciones gráficamente observando el digrafo de la figura 6-4.◀

Dominancia parcial

Existen relaciones en las que si un individuo domina a otro, éste no domina al primero, pero en las que no necesariamente en toda pareja del grupo existe una relación de dominio. Entonces se dice que en el grupo hay una relación de **dominancia parcial**. En este caso, si $A = [a_{ij}]$ es la matriz de incidencia se tiene: todos los a_{ij} son ceros o unos; $a_{ii} = 0$ para todo i; si $a_{ij} = 1$, entonces $a_{ji} = 0$. Como antes, la relación de dominancia parcial se puede transmitir a través de alguna k-cadena.

► **Ejemplo 6.11** Consideremos la relación representada en el digrafo contenido en la figura 6-5.

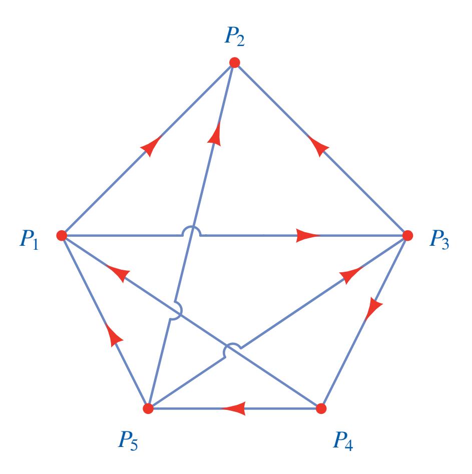


Figura 6-5 •

Entonces, la matriz de incidencia para éste es

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}.$$

Luego, la relación es de dominancia parcial, pues $a_{ij} = 1 \Rightarrow a_{ji} = 0$, las entradas en la diagonal son nulas y todas las componentes de esta matriz son unos o ceros.◀

► **Ejemplo 6.12** Si A es la matriz de incidencia de la relación de dominancia parcial del ejemplo anterior, entonces,

$$A + A^2 = \begin{bmatrix} 0 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 2 & 2 & 0 & 1 \\ 1 & 3 & 2 & 1 & 0 \end{bmatrix}.$$

De donde, por el corolario 6.1, se desprende que P_3, P_4 y P_5 dominan a todos los demás en, a lo sumo, dos etapas; mientras que P_2 es dominado por todos los demás, cuando mucho, dos etapas y P_2 nunca domina a individuo alguno en este grupo.◄

6.2 Redes de conducción y principios de conservación

Requisitos: *Sistemas lineales y método de Gauss: 1.2.1 a 1.2.4*

En una gran variedad de situaciones es posible resolver problemas utilizando principios básicos de conservación. En esta sección veremos unas cuantas aplicaciones del álgebra lineal que utilizan esta clase de principios de manera natural. La primera tiene que ver con redes muy simples que conducen cierto tipo de fluido, éste puede ser líquido, gas, electricidad, mensajes, vehículos, personas, etc. La figura 6-6 contiene un esquema de una de estas redes. Las líneas representan los conductores de la red, las flechas indican las direcciones hacia donde se mueve el fluido cuando cruza determinada sección de la red a través de un conductor; los puntos, llamados nodos, es donde convergen dos o más conductores. En los nodos se producen bifurcaciones donde el fluido se reparte a distintos conductores. Cada literal representa la magnitud del fluido que atraviesa el respectivo conductor en una sección de la red.

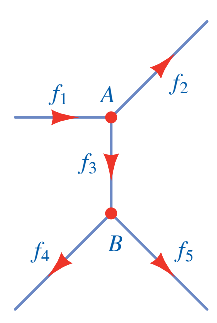


Figura 6-6 •

El germen que rige las redes conductoras es el siguiente:

Principio de los nodos: *En cada nodo, la suma de las magnitudes del fluido que entra a él, es igual a la suma de las magnitudes del fluido que sale del mismo.*

Así, en la red de la figura 6-6, se cumple:

- En el nodo A: $f_1 = f_2 + f_3$.
- En el nodo B: $f_3 = f_4 + f_5$.

Conociendo las magnitudes de algunos de los flujos y utilizando el principio de los nodos es posible, en general, establecer un sistema de ecuaciones lineales cuyo conjunto de soluciones determina los posibles flujos en la red conductora. Con frecuencia, las direcciones de los flujos que aparecen como variables en el esquema de una red conductora son convencionales y las direcciones dependen de los valores que resulten al resolver el sistema; entonces, si una de las soluciones f_i resulta negativa, significa que en realidad el flujo correspondiente tiene dirección opuesta a la que se le asignó en el diagrama; sin

embargo, en algunas redes, por ejemplo en circuitos vehiculares, las direcciones son fijas y no se pueden invertir. En tal caso, es necesario imponer restricciones a los parámetros del conjunto solución del sistema de ecuaciones para que los valores de las f_i sean no negativos.

La segunda aplicación trata un tipo particularmente importante de redes de conducción, circuitos o redes eléctricas que están conformados por mallas de conductores, resistores y baterías. Para analizar las redes eléctricas, además del principio de los nodos, se requiere de un principio físico de conservación que tiene que ver con los potenciales eléctricos en cada malla del circuito, contenido en la llamada segunda regla o segunda ley de Kirchhoff. Finalmente, la última aplicación –balance químico– que trataremos en este apartado, es consecuencia del principio básico de la naturaleza en el que se establece que la materia no se crea ni se destruye, sino que se transforma en otro tipo de materia o energía.

6.2.1 Flujo vehicular

Supongamos que en una zona de cierta ciudad el flujo vehicular se comporta como lo muestra la figura 6-7; donde las direcciones de las flechas indican las sentidos en cada avenida y las cantidades al lado, o por encima de las mismas, el volumen promedio de automóviles que atraviesan por hora cada sección. Se requieren determinar los volúmenes promedio en las intersecciones; esto es, los valores x_i, $i = 1, 2, 3$ y 4 de la figura 6-7. El problema lo podemos resolver utilizando el principio de los nodos, pues en cada una de las intersecciones la suma de los flujos de entrada debe ser igual a la suma de los flujos de salida; así:

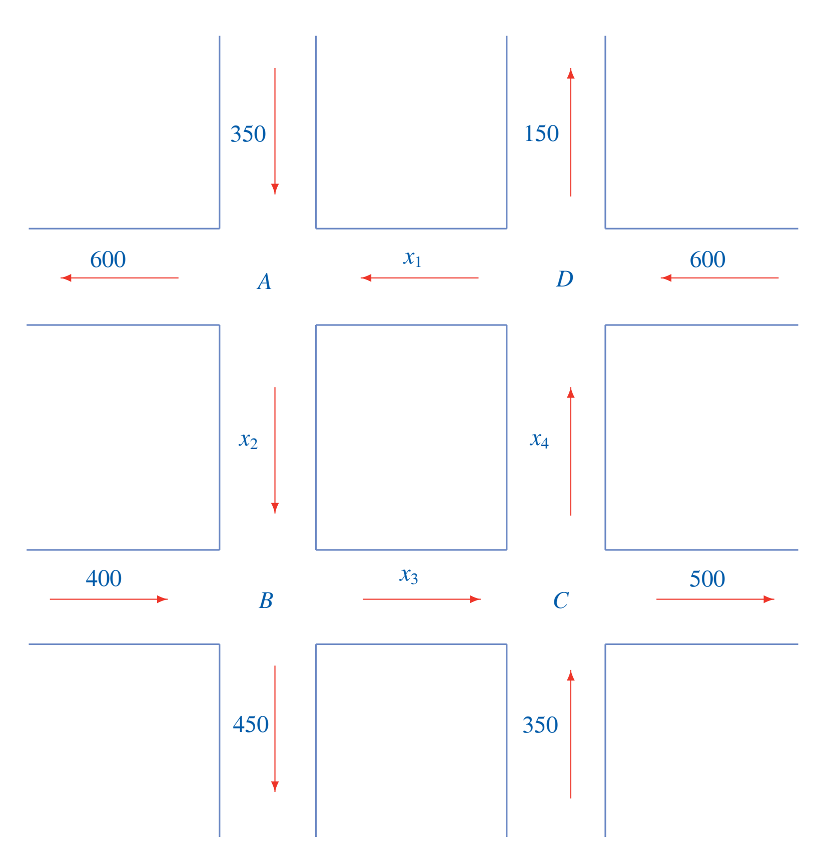


Figura 6-7 • Flujo vehicular en una zona de cierta ciudad.

En la intersección A: $x_1 + 350 = x_2 + 600$

En la intersección B: $x_2 + 400 = x_3 + 450$

En la intersección C: $x_3 + 350 = x_4 + 500$

En la intersección D: $x_4 + 600 = x_1 + 150$

Llevemos a forma escalonada la matriz aumentada del sistema,

$$\left[\begin{array}{rrrr|r} 1 & -1 & 0 & 0 & 250 \\ 0 & 1 & -1 & 0 & 50 \\ 0 & 0 & 1 & -1 & 150 \\ 1 & 0 & 0 & -1 & 450 \end{array}\right] \leftrightarrow \left[\begin{array}{rrrr|r} 1 & -1 & 0 & 0 & 250 \\ 0 & 1 & -1 & 0 & 50 \\ 0 & 0 & 1 & -1 & 150 \\ 0 & 1 & 0 & -1 & 200 \end{array}\right]$$

$$\leftrightarrow \left[\begin{array}{rrrr|r} 1 & -1 & 0 & 0 & 250 \\ 0 & 1 & -1 & 0 & 50 \\ 0 & 0 & 1 & -1 & 150 \\ 0 & 0 & 1 & -1 & 150 \end{array}\right]$$

$$\leftrightarrow \left[\begin{array}{rrrr|r} 1 & -1 & 0 & 0 & 250 \\ 0 & 1 & -1 & 0 & 50 \\ 0 & 0 & 1 & -1 & 150 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

y hagamos sustitución regresiva para obtener

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 450 + r \\ 200 + r \\ 150 + r \\ r \end{bmatrix} ;\, r \in \mathbb{R}.$$

Entonces, si se conoce el flujo vehicular entre C y D, se pueden determinar los demás flujos, por ejemplo, si $x_4 = 150$, entonces

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 600 \\ 350 \\ 300 \\ 150 \end{bmatrix}$$

es decir, el problema no tiene solución única. Sin embargo, sirve para modelar el tráfico vehicular cambiando el valor de x_4 (o especificando el valor de una, cualquiera, de las otras variables) para predecir los valores de las demás, siempre y cuando todas se conserven no negativas.

6.2.2 Circuitos eléctricos

En un circuito eléctrico es posible determinar la corriente en cada una de sus ramas en función de las resistencias y voltajes. Consideremos el circuito contenido en la figura 6-8. El símbolo —|⊢— representa una batería cuyo potencial eléctrico, o fuerza electromotriz ($fem = \varepsilon$), se mide en voltios (V); la batería produce una corriente eléctrica que fluye hacia afuera de la terminal indicada mediante la línea vertical más larga o, en forma esquemática, ⊂|⊢⊃. El resistor, representado por el símbolo —/\/\/\—, cuya

resistencia es medida en omhios (Ω), produce una caída en el potencial eléctrico gobernada por la ley de Ohm

$$E = IR,$$

donde I es la corriente, medida en amperios (A) y R la resistencia del resistor. Las flechas indican las direcciones de las corrientes en el circuito; sin embargo, si después de calcular alguna de ellas tiene signo negativo, significa que tiene dirección contraria a la que inicialmente se le asignó. Los nodos son los puntos donde se unen dos o más conductores en el circuito, en la figura 6-8 el dispositivo tiene dos nodos denotados por las letras A y B. Una malla en un circuito es cualquier recorrido conductor cerrado, en la figura 6-8 la trayectoria $ABCDA$ es una malla de esta red. Para resolver un circuito eléctrico se usan las llamadas leyes de Kirchhoff:

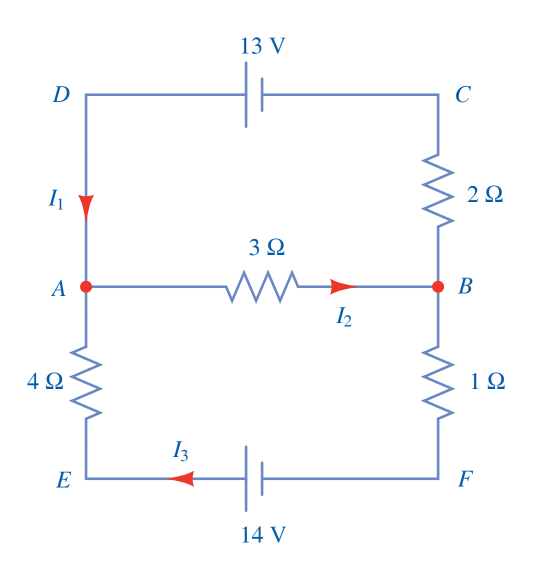


Figura 6-8 •

K1 En todo nodo la suma de las corrientes entrantes es igual a la suma de las corrientes salientes.

K2 En toda malla del circuito la suma algebraica de las fuerzas electromotrices es igual a la suma algebraica de las caídas de potencial en cada resistencia. Esto es, en una malla dada

$$\sum \varepsilon_j = \sum I_j R_j;$$

donde las ε_j son las fuerzas electromotrices en cada batería y los factores en los productos $I_j R_j$ están formados por las resistencias R_j de los resistores que están en la malla y las respectivas intensidades de corriente I_j que fluyen en cada uno de ellos.

Para resolver un circuito eléctrico es conveniente tener presentes los siguientes puntos:

1. Cuando se aplica la ley **K2** en una malla, se elige positivo un sentido de recorrido, por ejemplo, la dirección en sentido contrario al que avanzan las manecillas del reloj y negativo el sentido opuesto. Entonces, si una corriente está en contraflujo a la dirección elegida, se considera con signo negativo para calcular la caída de potencial en las resistencias que atraviesa. De manera similar, toda fuerza electromotriz cuya corriente de salida esté en sentido contrario a la dirección elegida como positiva, se considera negativa.

2. Si un circuito tiene n nodos, la ley **K1** se debe aplicar únicamente a cualquier subconjunto de $n-1$ nodos; pues la ecuación que se obtenga con esta ley en el enésimo nodo será redundante.
3. La ley **K2** se debe aplicar a cada una de las mallas que forman el circuito.

▶ **Ejemplo 6.13** Encontrar las corrientes I_j del circuito contenido en la figura 6-8.◀

Solución En virtud de **K1**, en el nodo A se tiene

$$I_1 + I_3 = I_2;$$

mientras que en el nodo B

$$I_2 = I_1 + I_3.$$

Por **K2**, en la malla $ABCDA$ (se eligió positivo el sentido contrario al que avanzan las manecillas del reloj),

$$2I_1 + 3I_2 = 13$$

y en la malla $ABFEA$ (se consideró positivo el sentido en el que avanzan las manecillas del reloj),

$$3I_2 + 5I_3 = 14.$$

De donde se obtiene el sistema (observe que las dos primeras ecuaciones son idénticas, por eso es recomendable tener en cuenta el punto 2 de las precedentes recomendaciones)

$$\begin{array}{rcrcrcr} I_1 & - & I_2 & + & I_3 & = & 0 \\ I_1 & - & I_2 & + & I_3 & = & 0 \\ 2I_1 & + & 3I_2 & & & = & 13 \\ & & 3I_2 & + & 5I_3 & = & 14 \end{array}$$

ya que

$$\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 0 \\ 2 & 3 & 0 & 13 \\ 0 & 3 & 5 & 14 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 2 & 3 & 0 & 13 \\ 0 & 3 & 5 & 14 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 5 & -2 & 13 \\ 0 & 3 & 5 & 14 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 5 & -2 & 13 \\ 0 & 0 & 31 & 31 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

al hacer sustitución regresiva resulta

$$I_3 = 1\text{A},\ I_2 = 3\text{A},\ I_1 = 2\text{A}. \quad ✔$$

▶ **Ejemplo 6.14** Encontrar las corrientes I_j del circuito contenido en la figura 6-9.◀

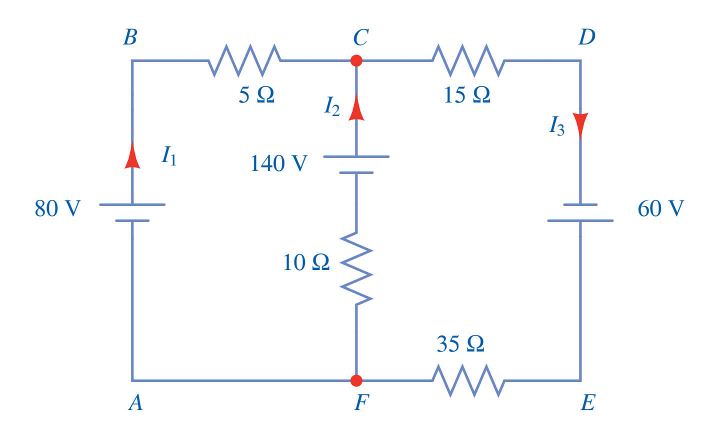


Figura 6-9 •

Solución Aplicando **K1** al nodo C se obtiene

$$I_1 + I_2 = I_3.$$

Aplicando **K2** a la malla $CDEFC$ (se considera positiva la dirección en el sentido en el que avanzan las manecillas del reloj),

$$10I_2 + 50I_3 = 140 + 60;$$

y aplicando **K2** ahora a la malla $ABCFA$ (se eligió la dirección positiva en el sentido en el que avanzan las manecillas del reloj),

$$5I_1 - 10I_2 = 80 - 140.$$

El sistema a resolver es entonces,

$$\begin{array}{rcrcrcr} I_1 & + & I_2 & - & I_3 & = & 0 \\ & & 10I_2 & + & 50I_3 & = & 200 \\ 5I_1 & - & 10I_2 & & & = & -60 \end{array}$$

Y, ya que

$$\left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 0 & 10 & 50 & 200 \\ 5 & -10 & 0 & -60 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 0 & 1 & 5 & 20 \\ 0 & -3 & 1 & -12 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 0 & 1 & 5 & 20 \\ 0 & 0 & 16 & 48 \end{array}\right],$$

al hacer sustitución regresiva resulta

$$I_1 = -2\,\text{A},\ I_2 = 5A,\ I_3 = 3\,\text{A};$$

donde el signo menos en $I_1 = -2\,\text{A}$, significa que la corriente I_1 fluye en sentido contrario al que le asignamos en la figura 6-9. ✔

6.2.3 Balance químico

Las reacciones químicas se representan simbólicamente mediante las llamadas *ecuaciones químicas*. Cuando moléculas con dos atómos de hidrógeno (H_2) arden, reaccionan con moléculas con 2 atómos de oxígeno (O_2) del aire, para formar moléculas de agua (H_2O). La ecuación química de esta reacción es

$$2H_2 + O_2 \longrightarrow 2H_2O \tag{6.1}$$

El signo + se lee como "reacciona con" y la flecha "produce". En el lado izquierdo de la ecuación química se encuentran las sustancias de partida, se les llama reactivas y en el lado derecho de la flecha las sustancias que se producen con la reacción, productos, separados por el símbolo + cuando hay más de una sustancia producto; los números que anteceden a los términos de la ecuación se llaman coeficientes y, de manera análoga al álgebra, no se escriben si son iguales a la unidad. Los subíndices significan el número de átomos de un elemento que forma una molécula; por ejemplo, H_2 representa una molécula con dos átomos de hidrógeno (otra vez, si no aparece un subíndice es que éste es uno). Los coeficientes indican el número de moléculas del término; así, en la ecuación (6.1) el coeficiente 2 en $2H_2$, indica que se tienen dos moléculas con dos átomos de hidrógeno cada una; mientras que el coeficiente 2 de H_2O en el lado derecho de (6.1), significa que se producen dos moléculas de agua con dos átomos de hidrógeno y uno de oxígeno cada una. Dado que en la naturaleza no se crea ni se destruye materia, **en toda ecuación química, para cualquiera de los elementos que la componen, el número de átomos de un elemento es igual en ambos lados** (de la flecha). De esta manera, en la ecuación (6.1) en el lado izquierdo se tienen $2 \times 2 = 4$ átomos de hidrógeno y 2 átomos de oxígeno; mientras que en el lado derecho 2(2 átomos de hidrógeno + un átomo de óxigeno) = 4 átomos de hidrógeno y 2 átomos de oxígeno. Cuando se cumple con el principio de conservación en una ecuación química, se dice que está **balanceada**. Sin embargo, en general se conocen las sustancias reactivas y las sustancias producto y se escribe la ecuación química sin balancear; entonces se tiene que encontrar los coeficientes, para que se cumpla el principio de conservación de la igualdad en la cantidad de átomos de cada elemento en ambos lados de la ecuación; a este proceso se le llama **balancear** la ecuación química. Por ejemplo, consideremos la siguiente reacción química

$$CH_4 + O_2 \longrightarrow CO_2 + H_2O$$

de las sustancias reactivas metano y óxigeno para producir dióxido de carbono y vapor de agua. Sean

x_1 : número de moléculas de CH_2
x_2 : número de moléculas de O_2

que tienen que reccionar para producir

x_3 : número de moléculas de CO_2
x_4 : número de moléculas de H_2O

de tal manera que la ecuación química

$$x_1CH_4 + x_2O_2 \longrightarrow x_3CO_2 + x_4H_2O$$

esté balanceada. Entonces, por el principio de conservación en la igualdad del número de átomos por cada elemento en ambos lados de la ecuación química, se debe tener:

- el número de átomos de carbono (C) debe ser igual en ambos lados: $x_1 - x_3 = 0$,
- el número de átomos de hidrógeno (H) debe ser igual en ambos lados: $4x_1 - 2x_4 = 0$,
- el número de átomos de óxigeno (O) debe ser igual en ambos lados: $2x_2 - 2x_3 - x_4 = 0$.

Que produce el sistema lineal homogéneo

$$\begin{array}{rcrcrcrcl} x_1 & & & - & x_3 & & & = & 0 \\ 4x_1 & & & & & - & 2x_4 & = & 0 \\ & & 2x_2 & - & 2x_3 & - & x_4 & = & 0 \end{array}$$

y ya que

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 4 & 0 & 0 & -2 \\ 0 & 2 & -2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 4 & -2 \\ 0 & 2 & -2 & -1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 2 & -2 & -1 \\ 0 & 0 & 4 & -2 \end{bmatrix};$$

al hacer sustitución regresiva se obtiene

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = r \begin{bmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \\ 1 \end{bmatrix}; r \in \mathbb{R}$$

como solución. Como es evidente que los x_i deben ser enteros no negativos, se puede tomar r igual a cualquier múltiplo entero positivo de 2; sin embargo, *se acostrumbra siempre elegir el menor entero para balancear la ecuación*. En este caso $r = 2$. Luego

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix}$$

es la solución buscada y la ecuación balanceada resultante es, entonces,

$$CH_4 + 2O_2 \longrightarrow CO_2 + 2H_2O.$$

6.3 Análisis insumo-producto

6.3.1 Modelo para economía abierta

Supongamos que una economía se divide en n industrias y cada una de ellas fabrica un solo producto final. Como es natural, cada una utiliza algunos productos (insumos) de las otras para fabricar sus propios artículos. Además, cada industria debe generar productos terminados para la demanda final. El análisis de insumo-producto tiene por objetivo determinar la producción de cada una de las industrias para satisfacer la demanda final si ésta cambia, suponiendo que la estructura económica no varía. La

Requisitos: *Sistemas lineales, matrices invertibles y método de Gauss-Jordan: 1.2.4, 1.2.5 y 2.1.1 a 2.1.3.*

Productor	Consumidor I_1	I_2	$\cdots$	I_n	Demanda final	Producción total
I_1	b_{11}	b_{12}	$\cdots$	b_{1n}	h_1	x_1
I_2	b_{21}	b_{22}	$\cdots$	b_{2n}	h_2	x_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
I_n	b_{n1}	b_{n2}	$\cdots$	b_{nn}	h_n	x_n

Tabla 6-1 • Tabla de insumo-demanda-producción.

tabla 6-1, es una **tabla** de **insumo-demanda-producción** que contiene la información necesaria para llevar a cabo este análisis. Cada elemento b_{ij} es el importe, en unidades monetarias, del producto que fabrica la industra I_i y requiere la industria I_j. Cada h_i es el importe, en unidades monetarias, que se demanda finalmente del producto que fabrica la industria I_i y cada x_i es la producción total necesaria, también en unidades monetarias, del producto fabricado por I_i. Entonces se debe tener

$$x_i = \sum_{j=1}^{n} b_{ij} + h_i \quad \forall i. \tag{6.2}$$

Definamos

$$a_{ij} = \frac{b_{ij}}{x_j}$$

es decir, a_{ij} es el valor monetario de la industria I_i que la industria I_j debe adquirir para producir una unidad monetaria de su propio producto. Se tiene entonces

$$\begin{aligned} x_i &= \sum_{j=1}^{n} b_{ij} + h_i \\ &= \sum_{j=1} a_{ij}x_j + h_i. \end{aligned}$$

De esta forma, si

$$A = [a_{ij}],$$

el sistema de ecuaciones (6.2) equivale a la ecuación matricial

$$A\vec{x} + \vec{h} = \vec{x} \tag{6.3}$$

donde $\vec{x} = (x_1, x_2, \ldots, x_n) = [\, x_1 \quad x_2 \quad \cdots \quad x_n \,]^t$ y $\vec{h} = (h_1, h_2, \ldots, h_n) = [\, h_1 \quad h_2 \quad \cdots \quad h_n \,]^t$. Evidentemente (6.3) equivale a

$$(I_n - A)\vec{x} = \vec{h} \tag{6.4}$$

donde I_n es la matriz identidad de orden n. La matriz $I_n - A$ se conoce como **matriz de Leontief**[1] (para econonomía abierta), por el creador de estas ideas y ganador del premio Nobel en economía en 1973 por sus estudios relacionados con el análisis insumo-producto (*input-output*) y a la matriz A se le llama

[1] Wassily Leontief (1906-1999).

matriz de insumos o **matriz de intercambio**. En el caso de que la matriz de Leontief sea invertible, la solución de estos sistemas equivalentes está dada por

$$\vec{x} = (I_n - A)^{-1}\vec{h}$$

siempre que[2] $\vec{x} \geq \vec{0}_{\mathbb{R}^n}$.

► **Ejemplo 6.15** Una economía hipotética consta de tres industrias I_1, I_2 e I_3, con la siguiente tabla de insumo-demanda-producción:

Productor	Consumidor I_1	I_2	I_3	Demanda final	Producción total
I_1	80	100	100	40	320
I_2	80	100	100	40	320
I_2	80	60	100	160	400

Determinar el correspondiente vector de producción $\vec{x}$ si la demanda final cambia a:

1. $\vec{h} = (84, 56, 35)$.
2. $\vec{h} = (56, 63, 70)$.

Las cifras están en millones de unidades monetarias.◄

Solución La matriz de insumos $A = [a_{ij}]$, $a_{ij} = \frac{b_{ij}}{x_j}$, está dada por

$$A = \begin{bmatrix} 80/320 & 100/320 & 100/400 \\ 80/320 & 100/320 & 100/400 \\ 80/320 & 60/320 & 100/400 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{5}{16} & \frac{1}{4} \\ \frac{1}{4} & \frac{5}{16} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{16} & \frac{1}{4} \end{bmatrix}$$

Entonces,

$$I_3 - A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{4} & \frac{5}{16} & \frac{1}{4} \\ \frac{1}{4} & \frac{5}{16} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{16} & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & -\frac{5}{16} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{11}{16} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{3}{16} & \frac{1}{4} \end{bmatrix}.$$

Calculemos, por el método de Gauss-Jordan, la inversa de la matriz de Leontief $I_3 - A$:

[2] Recuerde que si $U = [u_{ij}]$ y $V = [v_{ij}]$ son dos matrices del mismo tamaño, la notación $U < V$ significa $u_{ij} < v_{ij}$ para todo par ij; mientras $\vec{0}_{\mathbb{R}^n}$ representa el vector que tiene sus n componentes nulas, y que utilizamos las notaciones $\vec{u} = (u_1, u_2, \ldots, u_n)$ para representar indistintamente a esta n-eada ordenada, o a la matriz columna $[\ u_1 \quad u_2 \quad \cdots \quad u_n\ ]^t$.

$$\left[\begin{array}{rrr|rrr} \frac{3}{4} & -\frac{5}{16} & -\frac{1}{4} & 1 & 0 & 0 \\ -\frac{1}{4} & \frac{11}{16} & -\frac{1}{4} & 0 & 1 & 0 \\ -\frac{1}{4} & -\frac{3}{16} & \frac{3}{4} & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 1 & -\frac{5}{12} & -\frac{1}{3} & \frac{4}{3} & 0 & 0 \\ -\frac{1}{4} & \frac{11}{16} & -\frac{1}{4} & 0 & 1 & 0 \\ -\frac{1}{4} & -\frac{3}{16} & \frac{3}{4} & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|rrr} 1 & -\frac{5}{12} & -\frac{1}{3} & \frac{4}{3} & 0 & 0 \\ 0 & \frac{7}{12} & -\frac{1}{3} & \frac{1}{3} & 1 & 0 \\ 0 & -\frac{7}{24} & \frac{2}{3} & \frac{1}{3} & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|rrr} 1 & -\frac{5}{12} & -\frac{1}{3} & \frac{4}{3} & 0 & 0 \\ 0 & 1 & -\frac{4}{7} & \frac{4}{7} & \frac{12}{7} & 0 \\ 0 & -\frac{7}{24} & \frac{2}{3} & \frac{1}{3} & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|rrr} 1 & -\frac{5}{12} & -\frac{1}{3} & \frac{4}{3} & 0 & 0 \\ 0 & 1 & -\frac{4}{7} & \frac{4}{7} & \frac{12}{7} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|rrr} 1 & -\frac{5}{12} & -\frac{1}{3} & \frac{4}{3} & 0 & 0 \\ 0 & 1 & -\frac{4}{7} & \frac{4}{7} & \frac{12}{7} & 0 \\ 0 & 0 & 1 & 1 & 1 & 2 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|rrr} 1 & -\frac{5}{12} & 0 & \frac{5}{3} & \frac{1}{3} & \frac{2}{3} \\ 0 & 1 & 0 & \frac{8}{7} & \frac{16}{7} & \frac{8}{7} \\ 0 & 0 & 1 & 1 & 1 & 2 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & \frac{15}{7} & \frac{9}{7} & \frac{8}{7} \\ 0 & 1 & 0 & \frac{8}{7} & \frac{16}{7} & \frac{8}{7} \\ 0 & 0 & 1 & 1 & 1 & 2 \end{array}\right].$$

De donde

$$(I_3 - A)^{-1} = \begin{bmatrix} \frac{15}{7} & \frac{9}{7} & \frac{8}{7} \\ \frac{8}{7} & \frac{16}{7} & \frac{8}{7} \\ 1 & 1 & 2 \end{bmatrix}.$$

Por tanto:

1. Si $\vec{h} = (84, 56, 35)$, entonces

$$\begin{aligned} \vec{x} &= (I_3 - A)^{-1}\vec{h} \\ &= \begin{bmatrix} \frac{15}{7} & \frac{9}{7} & \frac{8}{7} \\ \frac{8}{7} & \frac{16}{7} & \frac{8}{7} \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 84 \\ 56 \\ 35 \end{bmatrix} \\ &= \begin{bmatrix} 292 \\ 264 \\ 210 \end{bmatrix}. \end{aligned}$$

2. Si $\vec{h} = (56, 63, 70)$,

$$\vec{x} = \begin{bmatrix} \frac{15}{7} & \frac{9}{7} & \frac{8}{7} \\ \frac{8}{7} & \frac{16}{7} & \frac{8}{7} \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 56 \\ 63 \\ 70 \end{bmatrix} = \begin{bmatrix} 281 \\ 288 \\ 259 \end{bmatrix}. \checkmark$$

Para resolver este tipo de problemas no siempre es necesario encontrar $(I_n - A)^{-1}$, donde A es la matriz de insumos, pues si únicamente se va a buscar el vector de producción para un solo vector de demanda dado $\vec{h}_1$, entonces es posible que convenga resolver por el método de Gauss el sistema $(I_n - A)\vec{x} = \vec{h}_1$, formando la matriz ampliada $[\ I_n - A \quad \vec{h}_1\ ]$ para llevarla a forma escalonada y hacer sustitución regresiva. Pero, dado que en la práctica la demanda final va cambiando por períodos, entonces el calcular una vez la inversa de la matriz de Leontief, y multiplicar por las correspondientes demandas finales, resulta más útil para estos casos. También, con frecuencia este tipo de problemas se plantea dando la información de la matriz de insumos, en vez de la tabla de insumo-demanda-producto, y la demanda final.

▶ **Ejemplo 6.16** Una economía está dividida en tres sectores: manufactura (I_1), agricultura (I_2) y servicios (I_3). La siguiente tabla contiene la información (hipotética) del intercambio de las cantidades de insumos entre ellas por unidad de producción:

Productor	Consumidor I_1	I_2	I_3
I_1	1/4	1/8	3/8
I_2	3/8	1/4	1/8
I_2	1/8	1/8	1/4

Encontrar las cantidades que tienen que producir (en millones de unidades monetarias) estas industrias para satisfacer la demanda final:

- 82 unidades del sector manufactura.
- 41 unidades del sector agricultura.
- 123 unidades del sector de servicios. ◀

Solución

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1/4 & 1/8 & 3/8 \\ 3/8 & 1/4 & 1/8 \\ 1/8 & 1/8 & 1/4 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & -\frac{1}{8} & -\frac{3}{8} \\ -\frac{3}{8} & \frac{3}{4} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{3}{4} \end{bmatrix}$$

Como el vector de demanda final es $\vec{h} = [\ 82 \quad 41 \quad 123\ ]V^t$, el vector de produción $\vec{x}$ es la solución del sistema

$$\left[\begin{array}{ccc|c} \frac{3}{4} & -\frac{1}{8} & -\frac{3}{8} & 82 \\ -\frac{3}{8} & \frac{3}{4} & -\frac{1}{8} & 41 \\ -\frac{1}{8} & -\frac{1}{8} & \frac{3}{4} & 123 \end{array}\right] \sim \left[\begin{array}{ccc|c} -1/8 & -1/8 & 3/4 & 123 \\ -3/8 & 3/4 & -1/8 & 41 \\ 3/4 & -1/8 & -3/8 & 82 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|c} 1 & 1 & -6 & -984 \\ -3/8 & 3/4 & -1/8 & 41 \\ 3/4 & -1/8 & -3/8 & 82 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|c} 1 & 1 & -6 & -984 \\ 0 & \frac{9}{8} & -\frac{19}{8} & -328 \\ 0 & -\frac{7}{8} & \frac{33}{8} & 820 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|c} 1 & 1 & -6 & -984 \\ 0 & 1 & -\frac{19}{9} & -\frac{2624}{9} \\ 0 & -\frac{7}{8} & \frac{33}{8} & 820 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|c} 1 & 1 & -6 & -984 \\ 0 & 1 & -\frac{19}{9} & -\frac{2624}{9} \\ 0 & 0 & \frac{41}{18} & \frac{5084}{9} \end{array}\right]$$

y al hacer sustitución regresiva obtenemos

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 272 \\ 232 \\ 248 \end{bmatrix}.$$

Es decir, el sector de manufactura debe producir 272 unidades, el sector agropecuario debe producir 232 unidades y el sector de servicios tiene que producir 248 unidades. ✔

Una cuestión básica, que ya mencionamos, es la posibilidad de que la matriz de Leontief pueda no ser invertible y si lo es, ello no garantiza $(I_n - A)^{-1}\vec{h} \geq \vec{0}_{\mathbb{R}^n}$ para todos los vectores de demanda admisibles. La suma de los elementos de una columna de la matriz de insumos representa la fracción de una unidad monetaria que la industria, que encabeza esa columna, gasta para crear una unidad monetaria de su propio producto. En una economía abierta existe un sector externo (exógeno), que consume la demanda final de cada sector industrial; este sector participa en el proceso, por ejemplo, con mano de obra; así que parte del gasto que hace cada industria debe destinarse a salarios, los cuales no están considerados en la matriz de insumos. Por tanto, en el modelo de Leontief, la suma de los elementos de cada columna debe ser inferior a una unidad monetaria; pues el resto se destina, por ejemplo, a pagar salarios. Esta propiedad es suficiente para garantizar la no singularidad de la matriz de Leontief y la no negatividad del producto de su inversa con los vectores de demanda que son admisibles. Enunciamos este resultado en la siguiente proposición y postergamos la demostración para el final de esta sección (apartado 6.3.4, proposición 6.9) donde también se incluye un método de aproximación para la inversa de la matriz de Leontief.

Proposición 6.1 *Sea A una matriz de insumos para una economía abierta; esto es, $A = [a_{ij}] \in \mathfrak{M}_n$, $a_{ij} \geq 0$ para todo par de subíndices i, j y $\sum_{i=1}^{n} a_{ij} < 1$, para todo $j = 1, 2, \ldots, n$. Entonces:*

1. *La matriz de Leontief $(I_n - A)$, donde I_n es la matriz identidad de orden n, es invertible.*
2. *Todas las componentes de la matriz $(I_n - A)^{-1}$ son no negativas.*
3. *$(I_n - A)^{-1}\vec{x} \geq \vec{0}_{\mathbb{R}^n}$ para todo $\vec{x} \geq \vec{0}_{\mathbb{R}^n}$.*

6.3.2 Modelo para economía cerrada

En el análisis insumo-producto supusimos que existe una demanda externa de los productos que fabrican los sectores industriales que forman el sistema económico productivo. Sin embargo, es posible que en economías más simples todo lo producido por esos sectores sea consumido en su totalidad por ellos mismos. En tal caso el modelo de Leontief, para economías cerradas, consiste en encontrar valores de equilibrio del sistema; es decir, la relación que deben tener las producciones de cada sector de tal manera que sus ingresos igualen a sus egresos. Supongamos que una economía sencilla tiene la siguiente tabla insumo-producción,

	Productor			
Consumidor	I_1	I_2	$\cdots$	I_n
I_1	a_{11}	a_{12}	$\cdots$	a_{1n}
I_2	a_{21}	a_{22}	$\cdots$	a_{2n}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
I_n	a_{n1}	a_{n2}	$\cdots$	a_{nn}

(6.5)

donde a_{ij} es la fracción del total de la producción de la industria I_j que requiere como insumo la industria I_i. En este modelo todo lo que produce una industria se consume en el sistema; así que los a_{ij} satisfacen $0 \leq a_{ij} \leq 1$ y $\sum_{i=1}^{n} a_{ij} = 1$ para todo j; es decir, la suma de los elementos en cada columna es uno. Nuevamente, a la matriz $A = [a_{ij}]$ se le llama matriz de insumos o matriz de intercambio. Sea $\vec{x} = (x_1, x_2, \ldots, x_n)$ el vector de producción; esto es, x_i es la cantidad, en unidades monetarias, que la industria I_i debe producir para que se cumplan las condiciones de equilibrio. Sea I_i una de estas industrias, entonces, de la tabla 6-5, el egreso total de I_i está dado por

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$$

puesto que el egreso debe ser igual al ingreso,

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = x_i\,.$$

Por tanto, se obtiene el sistema lineal homogéneo

$$\begin{array}{rcrcccrcl}
(1-a_{11})x_1 & - & a_{12}x_2 & - & \cdots & - & a_{1n}x_n & = & 0\\
-a_{21}x_1 & + & (1-a_{22})x_2 & - & \cdots & - & a_{2n}x_n & = & 0\\
 & \vdots & & \vdots & & \vdots & & \vdots & \\
-a_{n1}x_1 & - & a_{n2}x_2 & - & \cdots & - & (1-a_{nn})x_n & = & 0
\end{array}$$

Es decir,

$$(I_n - A)\vec{x} = \vec{0}_{\mathbb{R}^n} \qquad (6.6)$$

Llamaremos a la matriz $I_n - A$, matriz de Leontief para el modelo de economía cerrada.

► **Ejemplo 6.17** En una economía sencilla, la distribución insumo-producto se muestra en la siguiente tabla (hipotética):

Consumidor	Productor Agricultura	Construcción	Vestido
Agricultura	1/4	1/2	1/4
Construcción	1/2	1/4	1/4
Vestido	1/4	1/4	1/2

Encontrar qué condición debe cumplir el vector de producción para que los ingresos sean iguales a los egresos en cada sector◀.

Solución La matriz de insumo para este caso es

$$A = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$$

Resolvamos el sistema homogéneo (6.6):

$$\begin{aligned} I_3 - A &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix} \\ &= \begin{bmatrix} 3/4 & -1/2 & -1/4 \\ -1/2 & 3/4 & -1/4 \\ -1/4 & -1/4 & 1/2 \end{bmatrix} \\ &\sim \begin{bmatrix} -1/4 & -1/4 & 1/2 \\ -1/2 & 3/4 & -1/4 \\ 3/4 & -1/2 & -1/4 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & -2 \\ -1/2 & 3/4 & -1/4 \\ 3/4 & -1/2 & -1/4 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & -2 \\ 0 & 5/4 & -5/4 \\ 0 & -5/4 & 5/4 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & -2 \\ 0 & 5/4 & -5/4 \\ 0 & 0 & 0 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Al hacer sustitución regresiva obtenemos

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} r \\ r \\ r \end{bmatrix}, \; r > 0.$$

Es decir, los tres sectores deben producir el mismo número de unidades monetarias. ✔

Para que existan soluciones del sistema (6.6) es necesario que la matriz $I_n - A$ no sea invertible (contrariamente al modelo para economía abierta); sin embargo, como la economía es cerrada, todo se vende y todo se compra dentro del propio sistema; por tanto, la suma de las componentes de cada columna en la matriz de insumo debe resultar igual a 1. Esto garantizará que la matriz de Leontief para

el modelo de economía cerrada no sea invertible y que el sistema homogéneo (6.6) tenga una infinidad de soluciones. Debido a esto, en el modelo de economía cerrada no se resuelve el problema de cómo fijar los valores de la producción en el vector que las contiene, sino las relaciones numéricas que ellos tienen; en el ejemplo precedente la relación es $x_1 = x_2 = x_3$. El lector puede consultar la demostración de que la matriz de Leontief para el modelo de economía cerrada no es invertible en la proposición 6.2 del siguiente apartado dedicado a ese fin.

6.3.3 Singularidad de la matriz de Leontief para el modelo de economía cerrada

Requisitos: *Criterios de independencia lineal en $\mathbb{R}^n$: 3.3.1*

Recordemos que en el inciso *(b)*, del teorema 3.12, del apartado 3.3.1, mostramos que una matriz cuadrada de orden n no es invertible si y sólo si los n vectores que corresponden a sus columnas son linealmente dependientes. Es evidente, dado que una matriz es invertible si y sólo si su matriz transpuesta lo es, que la citada afirmación del teorema 3.12 sigue siendo válida si en lugar de considerar los vectores columna de la matriz, se aplica el criterio a los vectores fila de la matriz.

Proposición 6.2 *Sea $A = [a_{ij}]$ la matriz de insumos para el modelo de economía cerrada de Leontief, esto es:*

1. $0 \le a_{ij} \le 1$ *para todo* i, j.
2. $\sum_{i=1}^{n} a_{ij} = 1$ *para todo* $j = 1, 2, \ldots, n$.

Entonces, la matriz de Leontief para el modelo de economía cerrada, $I_n - A$, no es invertible.

DEMOSTRACIÓN ■

$$I_n - A = \begin{bmatrix} 1-a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ -a_{21} & 1-a_{22} & -a_{23} & \cdots & -a_{2n} \\ -a_{31} & -a_{32} & 1-a_{33} & \cdots & -a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & -a_{n3} & \cdots & 1-a_{nn} \end{bmatrix}. \tag{6.7}$$

y ya que para cada $j = 1, \ldots, n$, $\sum_{i=1}^{n} a_{ij} = 1$, entonces

$$\begin{aligned} 1-a_{11} &= -(-a_{21} - a_{31} - \cdots - a_{n1}) \\ -a_{12} &= -(1 - a_{22} - a_{32} - \cdots - a_{n2}) \\ -a_{13} &= -(-a_{23} + (1 - a_{33}) - \cdots - a_{n3}) \\ &\vdots \\ -a_{1n} &= -(-a_{2n} - a_{3n} - \cdots + (1 - a_{nn})); \end{aligned}$$

por tanto

$$\begin{aligned} (1-a_{11}, -a_{12}, -a_{13}, \ldots, -a_{1n}) &= (-1)(-a_{21}, 1-a_{22}, -a_{23}, \ldots, -a_{2n}) \\ &\quad + (-1)(-a_{31}, -a_{32}, 1-a_{33}, \ldots, -a_{3n}) \\ &\quad + \cdots + (-1)(-a_{n1}, -a_{n2}, -a_{n3}, \ldots, 1-a_{nn}). \end{aligned}$$

Por lo que las filas de la matriz son linealmente dependientes y, por ende, la matriz $I_n - A$ no es invertible. ■

6.3.4 Inversa de la matriz de Leontief para el modelo de economía abierta y método de aproximación

Requisitos: *Espacios vectoriales normados, nociones de límite de sucesiones reales: 4.2.*

Inversa de la matriz de Leontief

En el ejemplo 4.41 vimos que $\|(x_1,x_2,\ldots,x_n)\|_1 = |x_1| + |x_2| + \cdots + |x_n|$ es una norma en $\mathbb{R}^n$; en el ejemplo 4.43 probamos que si $A = [a_{ij}]$ es una matriz de tamaño $m \times n$, entonces $\|A\|_1 = \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|$, la suma de los valores absolutos de todas las componentes de la matriz y $\|A\|_\infty = \text{máx}\{\,|a_{ij}| : i = 1,\ldots,n, j = 1,\ldots,m\}$ son normas en el espacio vectorial $\mathfrak{M}_{m\times n}$. En el teorema 4.22 se hizo patente el hecho de que cualquier par de normas en un espacio vectorial de dimensión finita son equivalentes y que, por tanto, para fines de proximidad, se puede trabajar en cualquier espacio finitamente generado con la norma que más convenga. Vamos entonces a definir, en la siguiente proposición, una norma en el espacio de matrices cuadradas que será bastante cómoda para poder establecer la demostración de la proposición 6.1 y dar un método de aproximación a la inversa de la matriz de Leontief.

Proposición 6.3 *Sea $A = [a_{ij}] \in \mathfrak{M}_n$, con columnas $\vec{k}_j = (a_{1j}, a_{2j},\ldots,a_{nj}) = [\ a_{1j} \quad a_{2j} \quad \cdots \quad a_{nj}\ ]^t$, $j = 1,2,\ldots,n$; se define*

$$\|A\| = \max_{1\le j\le n} \|\vec{k}_j\|_1$$

es decir,

$$\|A\| = \max_{1\le j\le n} \left\{ \sum_{i=1}^{n} |a_{ij}| \right\}.$$

Entonces $\|\cdot\|$ es una norma en $\mathfrak{M}_n$.

DEMOSTRACIÓN ■ 1. Claramente $\|A\| \ge 0$ para toda matriz $A \in \mathfrak{M}_n$.

2. Si $A = \mathscr{O}$, evidentemente $\|A\| = 0$. Supongamos que $A = [a_{ij}] \in \mathfrak{M}_n$ y $\|A\| = 0$, entonces

$$|a_{ij}| \le \|A\| = 0$$

implica $A = \mathscr{O}$.

3. Sea $\lambda \in \mathbb{R}$ y $A = [a_{ij}] \in \mathfrak{M}_n$. Entonces

$$\begin{aligned}\|\lambda A\| &= \max_{1\le j\le n} \|\lambda\vec{k}_j\|_1 \\ &= \max_{1\le j\le n} |\lambda|\, \|\vec{k}_j\|_1 \\ &= |\lambda| \max_{1\le j\le n} \|\vec{k}_j\|_1 \\ &= |\lambda|\, \|A\|.\end{aligned}$$

4. Si $A = [a_{ij}]$ y $B = [b_{ij}]$ son matrices cuadradas de orden n con columnas $\vec{k}_j$ y $\vec{h}_j$, respectivamente, entonces

$$\begin{aligned}\|A+B\| &= \max_{1\le j\le n} \|\vec{k}_j+\vec{h}_j\|_1\\ &\le \max_{1\le j\le n}\left\{\|\vec{k}_j\|_1+\|\vec{h}_j\|_1\right\}\\ &\le \|A\|+\|B\|.\end{aligned}$$

De 1, 2, 3 y 4 se concluye que $\|\cdot\|$ es una norma en $\mathfrak{M}_n$. ■

La norma definida en la proposición anterior tiene una propiedad muy importante que no poseen otras normas, por ejemplo $\|\cdot\|_\infty$, (cfr. ejercicio 327 del capítulo 4); esta es la razón por la que hemos dicho que es una norma cómoda para alcanzar el propósito de este apartado. La propiedad a la que hacemos referencia está contenida en el segundo inciso de la siguiente proposición.

Proposición 6.4 *Sea $\|\cdot\|$ la norma definida en la proposición 6.3, entonces*

1. $$\|A\vec{x}\|_1 \le \|A\|\|\vec{x}\|_1$$
   para toda matriz $A\in\mathfrak{M}_n$ y para todo $\vec{x}\in\mathbb{R}^n$.

2. $$\|AB\| \le \|A\|\|B\|$$
   para todo par de matrices $A,B\in\mathfrak{M}_n$.

DEMOSTRACIÓN ■ 1. Sea $A\in\mathfrak{M}_n$, con columnas $\vec{k}_j$, y $\vec{x}=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$. Entonces (cfr. ejemplo 1.15)

$$\begin{aligned}\|A\vec{x}\|_1 &= \|x_1\vec{k}_1+x_2\vec{k}_2+\cdots+x_n\vec{k}_n\|\\ &\le |x_1|\|\vec{k}_1\|_1+|x_2|\|\vec{k}_2\|_1+\cdots+|x_n|\|\vec{k}_n\|_1\\ &\le |x_1|\|A\|+|x_2|\|A\|+\cdots+|x_n|\|A\|\\ &= \|A\|(|x_1|+|x_2|+\cdots+|x_n|)\\ &= \|A\|\|\vec{x}\|_1\,.\end{aligned}$$

2. Sean $A,B\in\mathfrak{M}_n$, con $\vec{h}_j$, $j=1,2,\ldots,n$, las columnas de B. Entonces (cfr. ejemplo 1.14)

$$AB=[\ A\vec{h}_1\quad A\vec{h}_2\quad\cdots\quad A\vec{h}_n\ ]$$

y, por tanto,

$$\begin{aligned}\|AB\| &= \max_{1\le j\le n}\|A\vec{h}_j\|_1\\ &\le \max_{1\le j\le n}\|A\|\|\vec{h}_j\|_1\\ &= \|A\|\|B\|.\ ■\end{aligned}$$

Definición 6.1 **(Convergencia de sucesiones de matrices)** *Sea $\{A_\nu\}$ una sucesión de matrices $A_1,A_2,\ldots,A_\nu,\ldots$ en el espacio de matrices cuadradas $\mathfrak{M}_n$. Se dice que la sucesión $\{A_\nu\}$ converge o tiene como límite a una matriz $L\in\mathfrak{M}_n$, cuando ν tiende a infinito y se escribe*

$$\lim_{\nu\to\infty}A_\nu=L,\quad \textit{si}\quad \lim_{\nu\to\infty}\|A_\nu-L\|=0.$$

Una sucesión de matrices cuadradas, $\{A_\nu\}$, converge a lo más a una matriz L. En efecto, si $L_1, L_2 \in \mathfrak{M}_n$ y la sucesión de matrices converge a L_1 y a L_2, entonces

$$\begin{aligned}\|L_1 - L_2\| &= \|(L_1 - A_\nu) + (A_\nu - L_2)\| \\ &\leq \|(L_1 - A_\nu)\| + \|(A_\nu - L_2)\|\end{aligned}$$

y como $\{A_\nu\}$ converge a L_1 y L_2, entonces $\|(L_1 - A_\nu)\|$, $\|(A_\nu - L_2)\|$ son arbitrariamente pequeños tomando ν suficientemente grande; lo cual implica

$$\|L_1 - L_2\| = 0$$

y, por tanto, $L_1 = L_2$.

❍ **Nota 6.1** Se puede probar, lo cual no haremos para no extendernos más, que una sucesión $\{A_\nu\}$ es convergente si y sólo si

$$\lim_{\nu,\mu\to\infty} \|A_\mu - A_\nu\| = 0.$$

Esta es un propiedad fundamental en las matemáticas que no se da únicamente en este caso, sino en cualquier espacio vectorial normado de dimensión finita. Pero no siempre es válida si la dimensión del espacio es infinita; de hecho, cuando se cumple en un espacio, se dice que éste es completo. Esta propiedad es de gran utilidad porque sirve para probar la convergencia de una sucesión aun cuando no sea posible exhibir explícitamente su límite.

Las siguientes dos propiedades son sencillas de probar y se dejan de ejercicio las demostraciones al lector.

Proposición 6.5 *Si* $\lim_{\nu\to\infty} A_\nu = L$ *y* $\lim_{\nu\to\infty} B_\nu = H$ *y* $\alpha \in \mathbb{R}$, *entonces:*

1. $$\lim_{\nu\to\infty} (A_\nu + B_\nu) = L + H.$$
2. $$\lim_{\nu\to\infty} \alpha A_\nu = \alpha L.$$

Proposición 6.6 *Sean* $\{A_\nu\}$ *una sucesión de matrices cuadradas que converge a* $L \in \mathfrak{M}_n$ *y* $M \in \mathfrak{M}_n$. *Entonces*

$$\lim_{\nu\to\infty} MA_\nu = ML.$$

DEMOSTRACIÓN ■

$$\begin{aligned}\lim_{\nu\to\infty} \|MA_\nu - ML\| &= \lim_{\nu\to\infty} \|M(A_\nu - L)\| \\ &\leq \lim_{\nu\to\infty} \|M\|\|A_\nu - L\| \\ &= \|M\| \lim_{\nu\to\infty} \|A_\nu - L\| \\ &= \|M\| \cdot 0 \\ &= 0;\end{aligned}$$

lo cual implica

$$\lim_{\nu\to\infty} MA_\nu = ML. \quad ■$$

Proposición 6.7 *Sea $A_\nu = [a_{ij}^{(\nu)}]$ una*[3] *sucesión de matrices cuadradas y $L = [l_{ij}] \in \mathfrak{M}_n$. Entonces*

$$\lim_{\nu\to\infty} A_\nu = L \Leftrightarrow \lim_{\nu\to\infty} a_{ij}^{(\nu)} = l_{ij} \ \ \forall\, i, j.$$

DEMOSTRACIÓN ■ ($\Rightarrow$) Supongamos que $\lim_{\nu\to\infty} A_\nu = L$, entonces si i, j son un par de índices fijos,

$$\begin{aligned}\left|a_{ij}^{(\nu)} - l_{ij}\right| &\le \sum_{k=1}^{n}\left|a_{kj}^{(\nu)} - l_{kj}\right| \\ &\le \|A_\nu - L\|\end{aligned}$$

de donde

$$\lim_{\nu\to\infty}\left|a_{ij}^{(\nu)} - l_{ij}\right| = 0.$$

($\Leftarrow$) Sean $\vec{k}_j^{(\nu)}$ las columnas de la matriz A_ν y $\vec{l}_j$ las columnas de la matriz L. Entonces $\lim_{\nu\to\infty} a_{ij}^{(\nu)} = l_{ij}$ $\forall\, i, j$ implica

$$\lim_{\nu\to\infty} \|\vec{k}_j^{(\nu)} - \vec{l}_j\|_1 = 0$$

y, por tanto,

$$\begin{aligned}\lim_{\nu\to\infty}\|A_\nu - L\| &= \lim_{\nu\to\infty}\max_{1\le j\le 1}\|\vec{k}_j^{(\nu)} - \vec{l}_j\|_1 \\ &= 0\,. \quad ■\end{aligned}$$

► **Ejemplo 6.18** Probar que si $A \in \mathfrak{M}_n$ es tal que $\|A\| < 1$, entonces

$$\lim_{\nu\to\infty} A^\nu = \mathscr{O}. ◄$$

DEMOSTRACIÓN ■ Sabemos que si $\{a_\nu\}$ es una sucesión de números reales, con $|a_\nu| < 1$ para todo ν, entonces $\lim_{\nu\to\infty} a_\nu = 0$; por tanto, de la proposición 6.4,

$$0 \le \lim_{\nu\to\infty}\|A^\nu\| \le \lim_{\nu\to\infty}\|A\|^\nu = 0,$$

luego

$$\lim_{\nu\to\infty}\|A^\nu\| = 0;$$

esto es,

$$\lim_{\nu\to\infty} A^\nu = \mathscr{O}\,. \quad ■$$

[3] En la notación $a_{ij}^{(\nu)}$, (ν) es un supraíndice, no un exponente.

Proposición 6.8 **(Serie geométrica en $\mathfrak{M}_n$)** *Sea $A \in \mathfrak{M}_n$ con $\|A\| < 1$. Se define $S_\nu = I_n + A + A^2 + \cdots + A^\nu$, para cada $v = 1, 2, \ldots,$ donde I_n es la matriz identidad de orden n. Entonces la sucesión $\{S_\nu\}$ es convergente.*

DEMOSTRACIÓN ■ Sabemos que una sucesión de números reales $\{a_n\}$ es convergente si y sólo si $\lim_{\nu,\mu\to\infty}(a_\nu - a_\mu) = 0$. Entonces, ya que, suponiendo sin perder generalidad que $\mu \geq \nu$,

$$\begin{aligned}\|S_\mu - S_\nu\| &= \|A^{\nu+1} + A^{\nu+2} + \cdots + A^\mu\| \\ &\leq \|A^{\nu+1}\| + \|A^{\nu+2}\| + \cdots + \|A^\mu\| \\ &\leq \|A\|^{\nu+1} + \|A\|^{\nu+2} + \cdots + \|A\|^\mu\end{aligned}$$

se desprende del hecho de que la serie geométrica $\sum_{k=1}^{\infty} \|A\|^K$ converge, pues $\|A\| < 1$, que

$$\begin{aligned}\lim_{\nu,\mu\to\infty} \|S_\mu - S_\nu\| &= \lim_{\nu,\mu\to\infty} \left(\|A\|^{\nu+1} + \|A\|^{\nu+2} + \cdots + \|A\|^\mu\right) \\ &= 0.\end{aligned}$$

Por tanto (cfr. nota 6.1), la sucesión $\{S_\nu\}$ es convergente. ■

Sea ahora $L = \lim_{\nu\to\infty} S_\nu$. Tenemos

$$S_\nu = I_n + A + A^2 + \cdots + A^\nu.$$

Multipliquemos ambos lados de la precedente igualdad por la matriz A para obtener

$$AS_\nu = A + A^2 + \cdots + A^\nu + A^{\nu+1};$$

entonces

$$S_\nu - AS_\nu = I_n - A^{\nu+1}$$

y, por tanto,

$$(I_n - A)S_\nu = I_n - A^{\nu+1} \tag{6.8}$$

Al tomar el límite cuando ν tiende a infinito (recuerde que $\|A\| < 1$), obtenemos

$$\begin{aligned}(I_n - A)L &= (I_n - A)\lim_{\nu\to\infty} S_\nu \\ &= \lim_{\nu\to\infty} (I_n - A)S_\nu \\ &= \lim_{\nu\to\infty} (I_n - A^{\nu+1}) \\ &= I_n - \lim_{\nu\to\infty} A^{\nu+1} \\ &= I_n - \mathcal{O} \\ &= I_n,\end{aligned}$$

lo cual prueba dos cosas:

- la matriz de Leontief $(I_n - A)$, $\|A\| < 1$, **es invertible**
- $L = (I_n - A)^{-1}$; esto es,

$$\begin{aligned}(I_n - A)^{-1} &= \lim_{\nu\to\infty} S_\nu \\ &= \lim_{\nu\to\infty} (I_n + A + A^2 + \cdots + A^\nu).\end{aligned}$$

Ahora bien, si A es un matriz de insumos, entonces las componentes de A son no negativas y la suma de ellas en cualquier columna de A es inferior a 1. Debido a que $S_\nu = I_n + A + A^2 + \cdots + A^\nu$ es suma de matrices con componentes no negativas, la matriz S_ν tiene componentes $\sigma_{ij}^{(\nu)} \geq 0$ para todo ν. Y ya que S_ν converge a $L = (I_n - A)^{-1}$, entonces cada componente $\sigma_{ij}^{(\nu)}$ converge (cfr. proposición 6.7) a la componente $(I_n - A)_{ij}^{-1}$ de la matriz $(I_n - A)^{-1}$; de donde se concluye que las componentes de la inversa de la matriz de Leontief son todas no negativas. De esta manera, si $\vec{x}$ es un vector de demanda, $(I_n - A)^{-1}\vec{x} \geq \vec{0}_{\mathbb{R}^n}$. Hemos probado así la proposición 6.1 que volvemos a hacer explícita aquí con el fin de resumir lo precedente.

Proposición 6.9 *Sea A una matriz de insumos para una economía abierta ($A = [a_{ij}] \in \mathfrak{M}_n$, $\|A\| < 1$ y $a_{ij} \geq 0$ para todo par de subíndices i, j). Entonces:*

1. *La matriz de Leontief $(I_n - A)$, donde I_n es la matriz identidad de orden n, es invertible.*
2. $$(I_n - A)^{-1} = \lim_{\nu\to\infty} (I_n + A + A^2 + \cdots + A^\nu).$$
3. *Todas las componentes de la matriz $(I_n - A)^{-1}$ son no negativas.*
4. $(I_n - A)^{-1}\vec{x} \geq \vec{0}_{\mathbb{R}^n}$ *para todo* $\vec{x} \geq \vec{0}_{\mathbb{R}^n}$.

Método de aproximación para la inversa de la matriz de Leontief

Es claro que calcular manualmente por el método de Gauss-Jordan la inversa de una matriz de tamaño grande puede ser una tarea sumamente difícil, o incluso imposible. Sin embargo, para el caso de la matriz de Leontief, podemos formular un método de aproximación que puede ser bastante útil. De la igualdad (6.8) se tiene

$$(I_n - A)(I_n + A + A^2 + \cdots + A^\nu) = I_n - A^{\nu+1}.$$

Entonces

$$(I_n - A)^{-1} \approx (I_n + A + A^2 + \cdots + A^\nu) \tag{6.9}$$

en la medida en la que

$$A^{\nu+1} \approx \mathscr{O}$$

de acuerdo con la norma $\|\cdot\|$; esto es, en la medida en la que

$$\|A\|^{\nu+1} \approx 0.$$

Puesto que $\|A\| < 1$, $\|A\|^{\nu+1}$ tiende a 0 cuando ν tiende a infinito. Entonces se requiere tener un criterio para decidir qué valor ν vamos a tomar para utilizar la aproximación (6.9); este valor es indudablemente aquel en el que se considere que por primera vez $\|A\|^{\nu+1}$ es muy cercana a cero. Para definir el significado de que $\|A\|^{\nu+1}$ sea cercana a cero, observemos que si T es un valor positivo cualquiera, entonces

$$\|A\|^{\nu+1} < T$$

implica que para todas las componentes de $A^{\nu+1}$ se tiene[4]

$$|a_{ij}^{(\nu+1)}| \leq \|A^{\nu+1}\| \leq \|A\|^{\nu+1} < T.$$

Entonces, si T es pequeño los valores absolutos de todas las componentes de $A^{\nu+1}$ son aún más pequeños. De esta manera el valor de T es un margen de tolerancia a partir del cual se fija el criterio para considerar que la matriz $A^{\nu+1}$ es suficientemente cercana a la matriz $\mathscr{O}$. Por ejemplo, si $T < 10^{-5} = 0.00001$, entonces $\|A\|^{\nu+1} < T$ implica que todos los valores absolutos de las componentes de $A^{\nu+1}$ son inferiores a 0.00001; entonces se podría utilizar la aproximación (6.9) con este criterio de tolerancia. Si ya se ha fijado una tolerancia T, podemos saber cuál es el valor ν más pequeño para el cual se tiene $\|A\|^{\nu+1} < T$, pues esta desigualdad equivale a

$$\nu > \frac{\ln(T)}{\ln(\|A\|)} - 1$$

y se elige así el primer entero que satisface esta desigualdad para obtener la aproximación

$$(I_n - A)^{-1} \approx I_n + A + A^2 + \cdots + A^{\nu}.$$

▶ **Ejemplo 6.19** Supongamos que

$$A = \begin{bmatrix} 0.01 & 0.02 & 0.01 \\ 0.02 & 0.00 & 0.03 \\ 0.00 & 0.00 & 0.01 \end{bmatrix}$$

es la matriz de insumos de un sistema abierto de economía y que la tolerancia fijada para la aproximación es $T = 0.001$. Entonces, la suma de las componentes de las columnas de A son: 0.03, 0.02 y 0.05, por tanto,

$$\|A\| = 0.05\,.$$

Luego,

$$\begin{aligned}\frac{\ln(T)}{\ln(\|A\|)} - 1 &= \frac{\ln(0.001)}{\ln(0.05)} - 1 \\ &\approx 2.0745\end{aligned}$$

De esta manera el primer entero ν que satisface

$$\nu > \frac{\ln(T)}{\ln(\|A\|)} - 1 \approx 2.0745$$

[4]En la notación $a_{ij}^{(\nu+1)}$, $(\nu + 1)$ es un supraíndice, no un exponente.

es

$$\nu = 3.$$

Así, la aproximación es (redondeando a 4 cifras decimales)

$$(I_3 - A)^{-1} \approx I_3 + A + A^2 + A^3$$

$$\approx \begin{bmatrix} 1.0105 & .02021 & 1.0819 \times 10^{-2} \\ 0.2021 & 1.0004 & 3.0519 \times 10^{-2} \\ 0 & 0 & 1.0101 \end{bmatrix}$$

Al hacer el producto de $(I_3 - A)$ con la aproximación obtenemos

$$(I_3 - A)\begin{bmatrix} 1.0105 & .02021 & 1.0819 \times 10^{-2} \\ 0.2021 & 1.0004 & 3.0519 \times 10^{-2} \\ 0 & 0 & 1.0101 \end{bmatrix} = \begin{bmatrix} .99999 & -1.0 \times 10^{-7} & -5.7 \times 10^{-7} \\ 0 & 1.0 & -3.8 \times 10^{-7} \\ 0 & 0 & 1.0 \end{bmatrix}$$

cuyo resultado es próximo a la matriz identidad.◀

▶ **Ejemplo 6.20** Sea la matriz de insumos

$$A = \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.1 \\ 0.1 & 0.2 & 0.3 & 0.2 \\ 0.2 & 0.2 & 0.1 & 0.2 \\ 0.1 & 0.2 & 0.1 & 0.1 \end{bmatrix}$$

Calcular el menor entero ν que se requiere para aproximar la inversa de matriz de Leontief para la matriz A con una tolerancia $T = 10^{-5}$.◀

Solución Las sumas de las componentes de cada columna son 0.5, 0.8, 0.8, y 0.6; por tanto, $\|A\| = 0.8$. Puesto que se requiere

$$\nu > \frac{\ln(10^{-5})}{\ln(.8)} - 1 = 50.594$$

la respuesta es $\nu = 51$. Así, la aproximación está dada por

$$(I_4 - A)^{-1} = I_4 + \sum_{k=1}^{51} A^k$$

y se requieren calcular las potencias sucesivas de la matriz A de 2 hasta 51. ✔

En la práctica, para sistemas grandes, o para sistemas con norma próxima a uno, es casi imposible hacer los cálculos manualmente para las potencias de A y se requiere de una computadora para realizar esta tarea y de métodos numéricos para calcular aproximadamente las potencias de la matriz de insumos (cfr. el capítulo 8).

6.4 Programación lineal

Requisitos: *Sistemas lineales y método de Gauss-Jordan: 1.2.1 a 1.2.5*

La programación lineal tiene como objetivo buscar máximos y/o mínimos de funciones lineales sujetas a restricciones también lineales. Aunque su base matemática es bastante elemental, fue hasta mediados del siglo anterior que surgieron problemas de optimización lineal. En tiempos de la Segunda Guerra Mundial, el ejército de Estados Unidos de América tuvo la necesidad de estudiar formas de asignar recursos materiales y humanos de la manera más eficiente y económica posible; fue entonces que los primeros problemas de programación lineal fueron resueltos. En la actualidad la optimización lineal se usa de manera sistemática en los negocios y en la industria para resolver problemas que podemos clasificar de *maximización de ganancias* y de *minimización de costos*. Delta Air Lines, por ejemplo, realiza a diario más de 2 500 escalas en sus vuelos nacionales, a Canadá y a México; empleando unas 450 aeronaves de 10 flotillas diferentes, cuyas capacidades, velocidades, niveles de ruido generado y otros parámetros pueden variar. El problema de asignación de flotillas consiste en hacer coincidir aeronaves con escalas de vuelo de modo que los asientos vayan ocupados. El patrón de vuelos de las aeronaves dentro del sistema de rutas se llama itinerario. Delta es una de las primeras líneas aéreas que ha resuelto por completo el problema de asignación de flotillas. El itinerario debe diseñarse para alcanzar la mayor ganancia posible, maximizando ingresos con el menor costo directo de operación permisible. Delta puso en práctica un algoritmo de programación lineal para asignar los distintos tipos de flotillas a las diferentes escalas de vuelo con el objetivo de minimizar los costos derivados sujetos a varias restricciones operativas. La restricción operativa más importante es la cantidad de aeronaves disponibles en cada flota. Otros factores también esenciales a considerar son: la planeación del mantenimiento, la correlación de habilidades de los pilotos con los distintos tipos de aeronaves, la asignación de tiempos de descanso para pilotos, alcance y velocidad de las aeronaves y restricciones de los aeropuertos. El problema de programación que Delta resuelve diariamente involucra unas 60 000 variables y 40 000 restricciones y se estima que estos programas han logrado ahorrar a esta línea aérea alrededor de 100 millones de dólares anuales.[5]

6.4.1 Enfoque geométrico

En esta sección introducimos el método gráfico de programación lineal cuando se tienen sólo dos variables. Para ello comenzamos con un problema de aplicación para motivar los conceptos básicos de programación lineal.

► **Ejemplo 6.21** Una compañía tiene dos talleres donde produce cuatro tipos de libreros para armar que se venden en tiendas de autoservicio. La siguiente tabla contiene la información relativa a los datos de producción, costos y demanda. Encontrar el número de días que tiene que trabajar cada taller durante seis meses para proporcionar, de la manera más económica, los libreros requeridos.

[5]Cfr. Radhka Subramanian, Richard Scheff, John Quillian, Steve Wiper, Roy Marsten, “Coldstart: Fleet Assignment at Delta Airlines”, Interfaces 24, núm. 1 (1994).

Tipo de librero	Producción diaria taller 1	Producción diaria taller 2	Demanda semestral
A	120	40	4 800
B	80	80	6 400
C	40	120	4 800
D	60	100	6 000
Costos diarios	\$40 000	\$60 000	

Planteamiento del problema: Sean x y y el número de días que tienen que trabajar los talleres 1 y 2, respectivamente. Dado que los costos diarios de producción son fijos para cada taller, la función de costo está dada por

$$C = 40\,000x + 60\,000y \tag{6.10}$$

Como se tienen que producir al menos 4 800 libreros del tipo A, 6 400 libreros del tipo B, 4 800 libreros del tipo C y 6 000 libreros del tipo D a un costo mínimo, lo que se requiere es:

Hallar el mínimo de $C = 40\,000x + 60\,000y$

$$\textit{sujeto a las restricciones} \quad \begin{aligned} 120x + 40y &\geq 4\,800 \\ 80x + 80y &\geq 6\,400 \\ 40x + 120y &\geq 4\,800 \\ 60x + 100y &\geq 6\,000 \\ x, y &\geq 0 \end{aligned} \tag{6.11}$$

Para resolver esta cuestión, necesitamos primero encontrar el conjunto de los pares (x, y) para los que se cumplen todas las restricciones y después buscar entre ellos a los que minimicen la función de costo. Es obvio que la solución (si existe) no es tan sencilla y se requiere de un análisis más profundo. Por eso estableceremos antes algunos tecnicismos y aspectos teóricos que serán la base para solucionar este tipo de problemas, llamados problemas de programación lineal.

Formato general de un problema de programación lineal de dos variables

Los problemas de programación lineal con dos variables se pueden plantear de la siguiente manera:

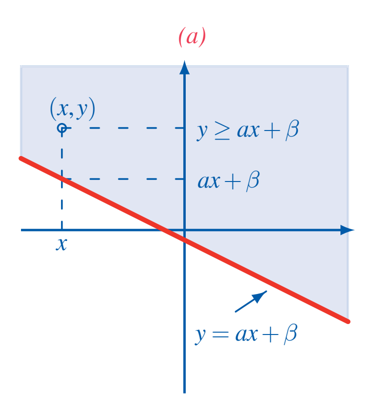
Maximizar (o minimizar) la función lineal

$$P = ax + by$$

sujeta a m restricciones lineales del tipo

$$a_{i1}x + a_{i2}y \begin{cases} \leq & d_i \quad \text{o} \\ = & d_i \quad \text{o} \\ \geq & d_i \end{cases} \tag{6.12}$$

donde se supone que x, y son variables no negativas. A la **función objetivo** $P = ax + by$ se le dice **función de ganancia** si lo que se pretende es maximizar y **función de costo** si lo que se desea es minimizar. Al conjunto solución de las restricciones (6.12) se le llama **conjunto factible** y a cualquier elemento del mismo **solución factible**. Así, en el ejemplo anterior, la función objetivo $C = 40\,000x + 60\,000y$ es una función de costo y todo punto que pertenece al conjunto solución de las restricciones (6.11) es una solución factible para este problema.



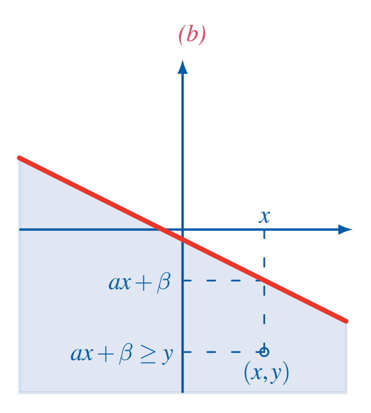


Figura 6-10 • *(a)* Semiplano $y \geq \alpha x + \beta$. *(b)* Semiplano $y \leq \alpha x + \beta$.

Semiplanos y conjuntos factibles

Al conjunto solución de una desigualdad del tipo $y \geq \alpha x + \beta$ se le denomina **semiplano** (cfr. ejercicio 417, pág. 412). Consiste en todos los puntos (x, y) del plano que están por encima y en la recta $y = \alpha x + \beta$ (cfr. figura 6-10*(a)*). De forma análoga se interpreta geométricamente el semiplano $y \leq \alpha x + \beta$; que es el conjunto de todos los pares (x, y) que están por debajo y en la recta $y = \alpha x + \beta$ (cfr. figura 6-10*(b)*). En todo problema de programación lineal las variables son no negativas, entonces, para encontrar el conjunto factible, se tienen que intersecar todos los conjuntos solución de cada una de las restricciones lineales (semiplanos) recordando que se deben limitar al primer cuadrante.

► **Ejemplo 6.22** Hallar el conjunto factible para las restricciones del ejemplo de los libreros (6.21).◄

Solución El conjunto solución de las restricciones lineales (6.11) equivale a la intersección, en el primer cuadrante de los semiplanos

$$\begin{aligned} y &\geq 120 - 3x \\ y &\geq 80 - x \\ y &\geq 40 - \tfrac{1}{3}x \\ y &\geq 60 - \tfrac{3}{5}x \end{aligned}$$

La figura 6-11 contiene un esbozo del conjunto factible de estas restricciones (sombreado). ✔

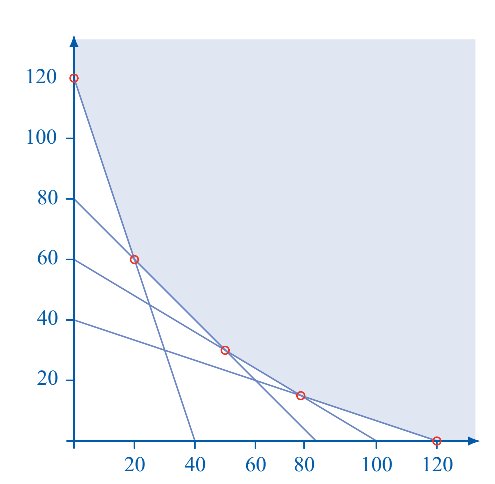
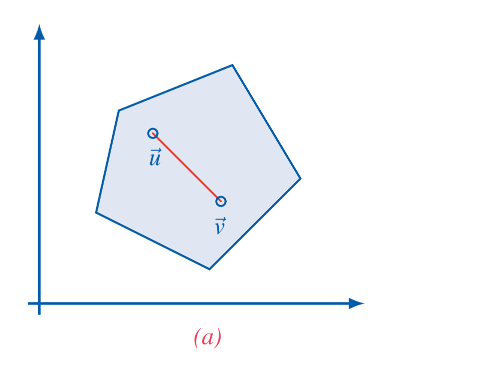


Figura 6-11 • Conjunto factible para el problema de la producción de libreros (ejemplos 6.21 y 6.22).



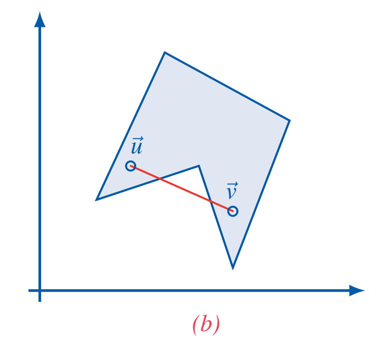


Figura 6-12 • *(a)* Conjunto convexo. *(b)* Conjunto no convexo.

Principio fundamental de la programación lineal

Un conjunto A del plano es **convexo** (cfr. ejercicio 416, pág. 412) si dado cualquier par de puntos $\vec{u}, \vec{v} \in A$, el segmento con punto inicial $\vec{u}$ y punto final $\vec{v}$:

$$[\vec{u}, \vec{v}] = \{\vec{u} + t(\vec{v} - \vec{u}) \,|\, 0 \leq t \leq 1\}$$

está completamente contenido en A. El conjunto de la figura 6-12*(a)* es convexo; mientras que el conjunto contenido en esta misma figura *(b)* no es convexo. No es difícil mostrar que si A y B son conjuntos convexos, entonces $A \cap B$ es un conjunto convexo y que todo semiplano (cfr. ejercicio 417, pág. 412) también es un conjunto convexo; luego, ya que todo conjunto factible es intersección finita de semiplanos, se desprende que **todo conjunto factible es un conjunto convexo** como el que se ilustra en la figura 6-13. Un segmento de recta formado por soluciones factibles que satisfacen $a_{i1}x + a_{i2}y = d_i$ para una de las restricciones (6.12) se denomina **arista** del conjunto factible. La intersección no vacía de dos o más aristas en un conjunto factible se llama **punto esquina** (o **vértice**) del mismo.

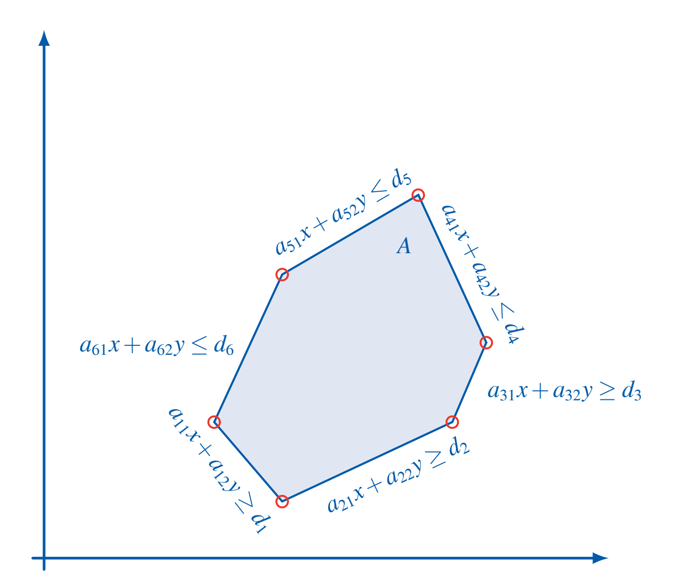


Figura 6-13 • Todo conjunto factible es un polígono convexo; a sus lados se les llama aristas y a sus vértices puntos esquina del conjunto factible.

Supongamos ahora que el conjunto convexo, contenido en la figura 6-13, es el conjunto factible para las restricciones

$$\begin{aligned} a_{11}x + a_{12}y &\geq d_1 \\ a_{21}x + a_{22}y &\geq d_2 \\ a_{31}x + a_{32}y &\geq d_3 \\ a_{41}x + a_{42}y &\leq d_4 \\ a_{51}x + a_{52}y &\leq d_5 \\ a_{61}x + a_{62}y &\leq d_6 \\ x, y &\geq 0 \end{aligned}$$

de un problema de programación lineal con función objetivo

$$f = ax + by,$$

y consideremos una solución factible (x_0, y_0). Para este punto la función objetivo tiene el valor

$$ax_0 + by_0 = f_0.$$

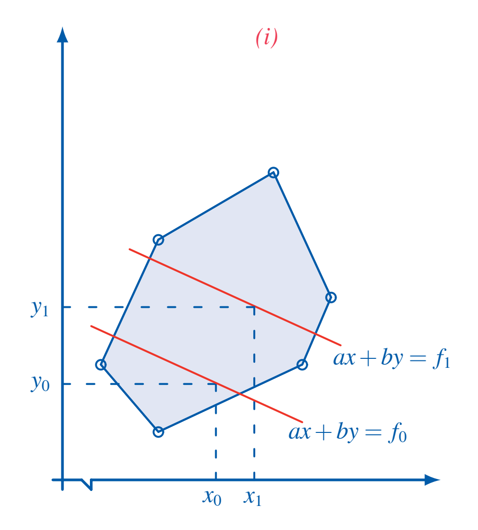
Por tanto, para todas las soluciones factibles (x, y) en la recta

$$ax + by = f_0$$

se obtiene el mismo valor f_0 al evaluar en la función objetivo. De manera análoga en todas las soluciones factibles de la recta

$$ax + by = f_1$$

la función objetivo tiene el valor constante f_1 y dado que las rectas $ax + by = f_0$ y $ax + by = f_1$ tienen la misma pendiente ($m = -a/b$), son rectas paralelas (cfr. figura 6-14(i)). Si suponemos (por simplicidad) que $b > 0$, el máximo de la función objetivo se alcanza en aquellos puntos del conjunto factible que pertenecen a la línea recta $ax + by = f$ que tiene mayor ordenada al origen y el mínimo corresponde a las soluciones factibles que están en la línea recta $ax + by = f$ que tiene menor ordenada al origen. De



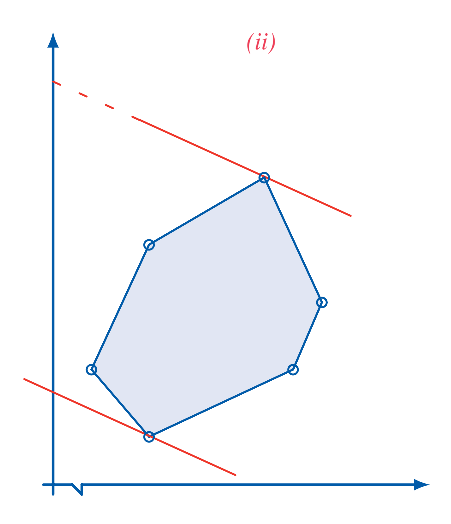


Figura 6-14 •

la figura 6-14*(ii)* se desprende que el máximo y el mínimo se alcanzan en puntos esquina del conjunto factible.[6] Con esta discusión hemos hecho plausible la demostración del siguiente teorema.

Teorema 6.3 (Principio fundamental de la programación lineal) *Si una función objetivo tiene un valor óptimo (máximo o mínimo), éste se alcanza en un punto esquina del conjunto factible.*

Entonces, para hallar el máximo (mínimo) de una función objetivo $f = ax + by$, se procede de la siguiente manera:

1. Graficar el conjunto factible definido por las restricciones.
2. Determinar si la función objetivo tiene máximo (mínimo).
3. De darse el paso 2, hallar las coordenadas de todos los puntos esquina del conjunto factible.
4. Evaluar la función objetivo $f = ax + by$ en cada punto esquina. El mayor (menor) valor obtenido será el máximo (mínimo) de la función objetivo en el conjunto factible.

► **Ejemplo 6.23 (Solución al ejemplo 6.21 de los libreros)** Ahora sí tenemos la herramienta adecuada para resolver el problema de los libreros (6.21) presentado al inicio de este segmento.◄

Solución Recordemos que el problema consiste en minimizar la función de costo

$$C = 40\,000x + 60\,000y$$

sujeta a las restricciones (6.11); es decir,

$$\begin{aligned} 120x + 40y &\geq 4\,800 \\ 80x + 80y &\geq 6\,400 \\ 40x + 120y &\geq 4\,800 \\ 60x + 100y &\geq 6\,000 \\ x, y &\geq 0 \end{aligned}$$

El esbozo del conjunto factible se hizo en el ejemplo 6.22 y está contenido en la figura 6-11 (cfr. pág. 615) Resolviendo las correspondientes ecuaciones vemos que los puntos esquina (vértices) del conjunto factible están dados por los puntos $(0,120)$, $(120,0)$, $(20,60)$ (intersección de las rectas $120x+40y = 4\,800$ y $80x+80y = 6\,400$), $(50,30)$ (intersección de las rectas $80x+80y = 6\,400$ y $60x+100y = 6\,000$), y $(75,15)$ (intersección de las rectas $60x+100y = 6\,000$ y $40x+120y = 4\,800$). Así, al evaluar la función de costo en cada uno de los puntos esquina tenemos:

Punto esquina	$C = 40\,000x + 60\,000y$
$(0,120)$	\$7 200 000
$(120,0)$	\$4 800 000
$(20,60)$	\$4 400 000
$(50,30)$	\$3 800 000
$(75,15)$	\$3 900 000

[6]Como es natural, es posible que el valor óptimo (máximo o mínimo) de una función objetivo se alcance en más de un punto del conjunto factible.

De donde se obtiene un costo mínimo de \$3 800 000 trabajando 50 días el taller 1 y 30 días el taller 2. Note que con esta solución óptima, trabajando 50 días el taller 1 y 30 el taller 2, se satisfacen exactamente las demandas de los libreros del tipo B y D (6 400 y 6 000 libreros, respectivamente); pero en el caso de los libreros del tipo A y C no sólo se cumple con la demanda (4 800), sino que además hay un excedente de 2 400 y 800 libreros, respectivamente, con un costo mínimo. ✔

Problemas no acotados

No todo problema de programación lineal alcanza un valor óptimo, cuando sucede así se dice que dicho problema **no es acotado**.

▶ **Ejemplo 6.24** Sean las funciones objetivo *(a)* $P = 4x + 2y$; *(b)* $P = -3x + 5y$; *(c)* $P = -2x - y$, y consideremos para cada una de estas funciones las restricciones

$$\begin{aligned} x &\geq 2 \\ y &\geq 3 \\ x, y &\geq 0 \end{aligned}$$

cuyo conjunto solución, S, está bosquejado en la figura 6-15. Entonces:

(a) Puesto que todo punto de la forma $(2, y)$, $y \geq 3$, es una solución factible y $\text{lím}_{y\to\infty}(4(2) + 2y) = \infty$, se desprende que $P = 4x + 2y$ no alcanza un valor máximo en el conjunto factible S. Claramente $P = 4x + 3y$ alcanza el valor mínimo 14 en $(2, 3) \in S$. Así, el problema de maximizar $P = 4x + 2y$ sujeto a las restricciones dadas no es acotado.

(b) Ya que todo punto de la forma $(x, 3)$, $x \geq 2$, pertenece al conjunto factible y $\text{lím}_{x\to\infty}(-3x + 5(3)) = -\infty$, se deduce que $P = -3x + 5y$ no alcanza un mínimo en S. De manera análoga, todo punto de la forma $(2, y)$, $y \geq 3$, pertenece al conjunto factible y $\text{lím}_{y\to\infty}(-3(2) + 5y) = \infty$; por tanto, $P = -3x + 5y$ no alcanza máximo en S. Entonces, los problemas de maximizar y minimizar $P = -3x + 5y$ sujetos a las restricciones dadas no son acotados.

(c) $(x, 3) \in S$ para todo $x \geq 2$ y $\text{lím}_{x\to\infty}(-2x - 3) = -\infty$ implican que $P = -2x - y$ no alcanza mínimo en S; sin embargo, $\text{máx}_{(x,y)\in S}\{-2x - y\} = -2(2) - 3 = -7$; es decir, $P = -2x - y$ alcanza su máximo, -7, en $(2, 3) \in S$. ◀

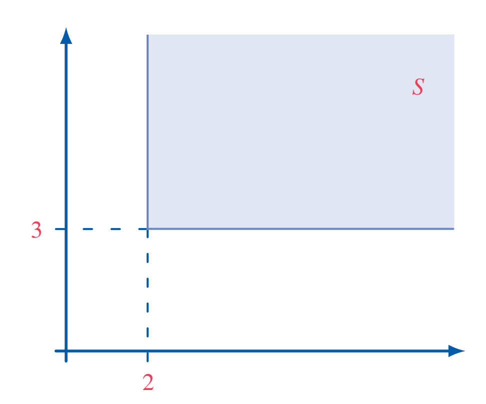


Figura 6-15 •

6.4.2 Método simplex para el problema estándar de programación lineal

Como se mencionó al inicio de esta sección, un problema de programación lineal puede involucrar más de dos variables y varias restricciones. El planteamiento general está dado de la siguiente manera:

$$\begin{array}{ll} \text{Optimizar (maximizar o minimizar)} & f = c_1x_1 + cx_2 + \cdots + c_nx_n \\ \text{sujeto a } m \text{ restricciones del tipo} & a_{i1}x_1 + a_{i2}x_2 + \cdots + a_ix_n \begin{cases} \leq b_i & \text{o} \\ = b_i & \text{o} \\ \geq b_i \end{cases} \\ & x_j \geq 0 \quad \forall j \end{array} \tag{6.13}$$

Nuevamente, el conjunto solución de las restricciones lineales (6.13) es el **conjunto factible** del problema de programación lineal y resulta ser un conjunto convexo; a cualquier elemento del conjunto factible se le llama **solución factible**; a todo subconjunto de un hiperplano[7] formado por soluciones factibles que satisfacen la ecuación $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_ix_n = b_i$, para alguna de las restricciones (6.13), se le dice **arista** o **borde** del conjunto factible, y a la intersección no vacía de n aristas se le denomina **punto esquina** del conjunto factible. Aunque el número de variables sea mayor a dos, **sigue siendo válido el principio fundamental de la programación lineal enunciado en el teorema** 6.3. Entonces, para resolver un problema de programación lineal con cualquier número de variables y restricciones, se tienen que encontrar todos los puntos esquina, evaluar la función objetivo y hallar así el valor óptimo (si es que la función objetivo alcanza uno). Este procedimiento puede implicar, en función del número de restricciones y variables, la resolución de un sistema considerablemente grande de ecuaciones lineales para encontrar los puntos esquina; lo cual puede ser, en la práctica, imposible aun con el uso de computadoras. En la década de los cuarenta, el matemático George Dantzig creó un ingenioso método para minimizar estos cálculos llamado **método simplex** y a su estudio está dedicado este apartado.

El problema estándar de programación lineal

Antes de estudiar el método simplex para restricciones generales del tipo (6.13) comenzamos con este algoritmo para el **problema estándar de programación lineal**; el cual se establece de la siguiente forma:

$$\begin{array}{ll} \text{Maximizar} & P = c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{sujeto a} & \left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & \leq & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & \leq & b_2 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & \leq & b_m \end{array} \right. \\ & x_j \geq 0, b_i \geq 0 \quad \forall\, i, j \end{array} \tag{6.14}$$

Si $\vec{c} = [\ c_1 \quad c_2 \quad \cdots \quad c_n\ ]^t, \vec{x} = [\ x_1 \quad x_2 \quad \cdots \quad x_n\ ]^t, \vec{b} = [\ b_1 \quad b_2 \quad \cdots \quad b_m\ ]^t$ y $A = [a_{ij}]$, $1 \leq i \leq m$, $1 \leq j \leq n$, el problema estándar de programación lineal (6.14) se puede escribir matricialmente como:

$$\begin{array}{ll} \text{Maximizar} & P = \vec{c}^{\,t}\vec{x} \\ \text{sujeto a} & A\vec{x} \leq \vec{b} \\ & \vec{x} \geq \vec{0}_{\mathbb{R}^n} \text{ y } \vec{b} \geq \vec{0}_{\mathbb{R}^m} \end{array} \tag{6.15}$$

[7]Cfr. ejercicio resuelto 9 del capítulo 5, pág. 502.

donde convenimos que para dos matrices $U = [u_{ij}]$, $V = [v_{ij}]$ del mismo tamaño, la notación $U \leq V$ significa $u_{ij} \leq v_{ij}$ para todo i, j. Cabe notar que[8] $\vec{x} = \vec{0}_{\mathbb{R}^n}$ **es siempre punto esquina de todo problema estándar de programación lineal,** ya que $\vec{b} \geq \vec{0}_{\mathbb{R}^m}$. Con el fin de motivar el método simplex para el problema estándar (6.14), ilustraremos las ideas fundamentales a través de un problema de dos variables contenido en el siguiente ejemplo.

► **Ejemplo 6.25** Una fábrica manufactura dos artículos; para producir éstos se utilizan tres máquinas. La primera puede usarse a lo sumo 80 horas, la segunda a lo más 50 horas y la tercera se puede utilizar un máximo de 90 horas. El primer artículo usa dos horas en la primera máquina y una hora en cada una de las otras dos máquinas; mientras que el segundo artículo requiere una hora en la primera, una hora en la segunda máquina y dos horas en la tercera. Si la utilidad es de \$500 por el primer artículo y \$700 por el segundo, ¿cuántas unidades de cada artículo deben fabricarse para maximizar la ganancia?◄

Solución Sean x_1 y x_2 el número de unidades del primero y del segundo artículo, respectivamente, que la fábrica manufactura. Entonces la función de ganancia está dada por

$$P = 500x_1 + 700x_2.$$

Puesto que la primera máquina se puede utilizar a lo más 80 horas, el primer artículo requiere 2 horas y el segundo 1 hora en esta máquina, se debe tener $2x_1 + x_2 \leq 80$. Ya que el primer artículo requiere 1 hora y el segundo 1 hora en la segunda máquina y ésta se puede utilizar a lo sumo 50 horas, se debe cumplir $x_1 + x_2 \leq 50$. Por último, como el primer artículo necesita 1 hora y el segundo 2 horas en la tercera máquina y ésta se puede emplear cuando mucho 90 horas, entonces $x_1 + 2x_2 \leq 90$. Así, el problema a resolver es:

$$\text{maximizar} \quad P = 500x_1 + 700x_2 \tag{6.16}$$

$$\text{sujeto a} \quad \left\{\begin{array}{rcrcr} 2x_1 & + & x_2 & \leq & 80 \\ x_1 & + & x_2 & \leq & 50 \\ x_1 & + & 2x_2 & \leq & 90 \end{array}\right. \tag{6.17}$$
$$x_1, x_2 \geq 0$$

La figura 6-16*(a)* contiene el conjunto factible, S, para este problema y los puntos esquina son $(0,0)$, $(0,45)$, $(10,40)$ (intersección de las rectas $x_1 + 2x_2 = 90$, $x_1 + x_2 = 50$), $(30,20)$ (intersección de las rectas $2x_1 + x_2 = 80$ y $x_1 + x_2 = 50$) y $(40,0)$. La evaluación de la función de ganancia en cada uno de los puntos esquina se resume en la siguiente tabla:

Punto esquina	$P = 500x_1 + 700x_2$
$(0,0)$	$P(0,0) = 0$
$(0,45)$	$P(0,45) = 31\,500$
$(10,40)$	$P(10,40) = 33\,000$
$(30,20)$	$P(30,20) = 29\,000$
$(40,0)$	$P(40,0) = 20\,000$

En la tabla se observa que el máximo se alcanza en $(10,40)$; es decir, se deben producir 10 unidades del primer artículo y 40 unidades del segundo para una ganancia de \$33 000. ✔

Resolvamos ahora el ejemplo precedente (6.25) con una perspectiva distinta, la del método simplex que, en esencia, consiste en tomar "atajos" a través de las aristas del conjunto factible hacia donde

[8] La notación $\vec{0}_{\mathbb{R}^k}$ representa el vector columna $k \times 1$: $\vec{0}_{\mathbb{R}^k} = [\ 0 \quad 0 \quad \cdots \quad 0\ ]^t$. para cada $k = 1, 2, \ldots$

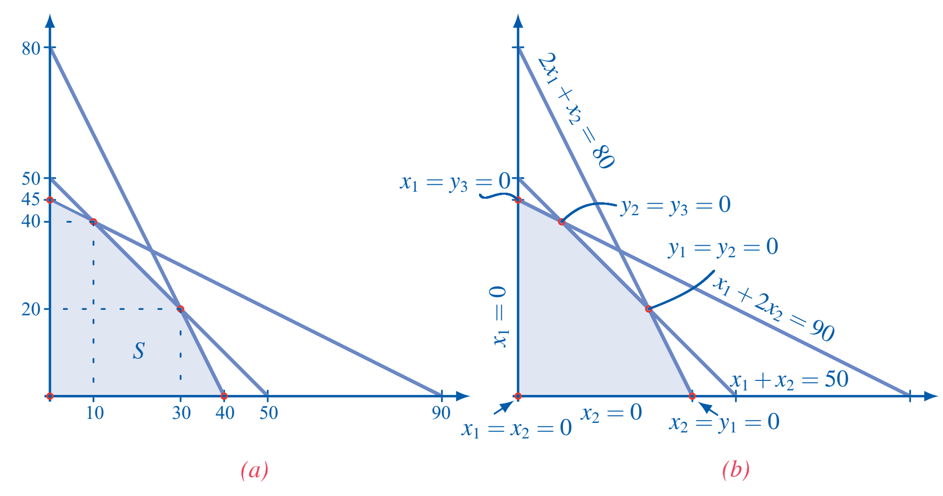


Figura 6-16 • *(a)* Conjunto factible y puntos esquina del ejemplo 6.25. *(b)* Si hacen ceros dos variables del sistema 6.18, la solución que resulta es un punto esquina del conjunto factible de las restricciones 6.17.

se produzca un incremento con mayor rapidez en la función de ganancia. Sea (x_1, x_2) cualquier punto del conjunto factible contenido en la figura 6-16, entonces se cumplen las restricciones (6.17); sean $y_1, y_2, y_3 \geq 0$ tales que

$$\begin{aligned} 2x_1 + x_2 + y_1 &= 80, \\ x_1 + x_2 + y_2 &= 50 \quad \text{y} \\ x_1 + 2x_2 + y_3 &= 90. \end{aligned}$$

Por ejemplo, si $(x_1, x_2) = (10, 20)$, entonces $y_1 = 40$, $y_2 = 20$ y $y_3 = 40$. De esta forma a cada solución factible (x_1, x_2) de este problema le corresponde una solución $(x_1, x_2, y_1, y_2, y_3)$ del sistema lineal

$$\begin{array}{rcrcrcrcrcl} 2x_1 & + & x_2 & + & y_1 & & & & & = & 80 \\ x_1 & + & x_2 & & & + & y_2 & & & = & 50 \\ x_1 & + & 2x_2 & & & & & + & y_3 & = & 90 \end{array} \qquad (6.18)$$

a la cual le llamaremos también solución factible y viceversa. Los valores y_1, y_2 y y_3 se llaman **variables de holgura** y para el problema en cuestión representan, respectivamente, el número de horas restantes disponibles para la primera, segunda y tercera máquina, una vez que se ha decidido fabricar x_1 unidades del primer artículo y x_2 unidades del segundo artículo. Por medio del sistema (6.18) es posible identificar los puntos esquina y aristas del conjunto factible; por ejemplo, el punto esquina $(0, 45)$ corresponde a la solución $(x_1, x_2, y_1, y_2, y_3)$ de (6.18) para $x_1 = 0$ y $y_3 = 0$ (por ejemplo, la solución factible $(0, 45, 35, 5, 0)$ del sistema (6.18)); el punto esquina (10, 40) corresponde a la solución $(x_1, x_2, y_1, y_2, y_3)$ de (6.18) para $y_2 = 0$ y $y_3 = 0$ (esto es, la solución factible (10, 40, 20, 0, 0) de (6.18)); etc. Esta información se sintetiza en la figura 6-16*(b)*. Podemos resumir la información total de las variables cuando (x_1, x_2) es el punto esquina (0, 0) en la siguiente tabla:

	x_1	x_2	y_1	y_2	y_3	
y_1	2	1	1	0	0	80
y_2	1	1	0	1	0	50
y_3	1	2	0	0	1	90
P	−500	−700	0	0	0	0

(6.19)

La primera fila contiene las variables del problema x_1, x_2 y las variables de holgura y_i; la primera columna contiene las variables de holgura; la última fila los coeficientes de la ecuación de costo (6.16)

pero en la forma $P - 500x_1 - 700x_2 - 0y_1 - 0y_2 - 0y_3 = 0$ (a esta fila se le llama **fila objetivo**); de las filas segunda a tercera y de las columnas segunda a última se ha escrito la matriz aumentada del sistema lineal (6.18). Esta se llama **tabla inicial** para el problema de programación inicial, ella contiene toda la información necesaria para saber el valor de todas las variables cuando (x_1, x_2) es el punto esquina $(0, 0)$:

- Obviamente, $x_1 = 0$ y $x_2 = 0$ porque hemos supuesto que (x_1, x_2) es el punto esquina $(0, 0)$; sin embargo, en la tabla inicial esta información se ve plasmada por el hecho de que ninguno de los vectores debajo de las etiquetas x_1 y x_2 de la primera fila es alguno de los vectores canónicos unitarios $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, o $(0, 0, 1, 0)$.
- Las variables de holgura, contenidas en la primera columna, toman los valores correspondientes en la última columna: $y_1 = 80$, $y_2 = 50$ y $y_3 = 90$. En este estado, $(x_1, x_2) = (0, 0)$, las tres variables de holgura son no nulas y este hecho se refleja en la tabla inicial al estar debajo de las etiquetas y_1, y_2 y y_3, de la primera fila, los vectores canónicos unitarios $(1, 0, 0, 0)$, $(0, 1, 0, 0)$ y $(0, 0, 1, 0)$, respectivamente.
- La función de ganancia P, en la primera columna, toma el valor del registro de la fila objetivo en la última columna: $P = 0$.

De la solución del ejemplo 6.25 sabemos que el máximo de este problema de producción se alcanza en el punto esquina $(10, 40)$. De la figura 6-16 es claro que podemos evitar la evaluación de la función de ganancia en los puntos esquina $(40, 0)$ y $(30, 20)$ si se toma un "atajo" al "avanzar" por la arista $x_1 = 0$ para llegar al punto esquina $(45, 0)$. ¿Cómo se puede obtener esta misma información de la tabla inicial? Al estar en el punto esquina $(0, 0)$ únicamente existen dos aristas a través de los cuales pueden incrementarse x_1 o x_2, que son $x_1 = 0$ o $x_2 = 0$ y dado que la función de ganancia es $P = 500x_1 + 700x_2$, entonces P tiene un incremento más rápido si se permite que x_2 sea distinto de cero (en este caso $P = 700x_2$, ya que $x_1 = 0$); es decir, al variar x_2 a lo largo de la arista $x_1 = 0$ (cfr. figura 6-16). Esta información se puede conseguir en la tabla inicial (6.19) eligiendo la variable que se encuentra en la columna que corresponde al registro negativo de mayor valor absoluto en la fila objetivo de la tabla 6.19 (la última fila). En nuestro caso concreto -700 es el registro con mayor valor absoluto y la variable que se encuentra arriba de este registro, en la misma columna, es x_2. A esta columna le llamaremos **columna pivote** para la tabla inicial. Al desplazarse a lo largo de la arista $x_1 = 0$ se llegará al punto esquina $(0, 45)$; nuevamente esta información la podemos obtener de la tabla inicial. Como el siguiente punto esquina a lo largo de la arista $x_1 = 0$ pertenece a cierta arista, entonces alguna de las variables de holgura, y_i, tiene que ser cero en el sistema (6.18). Podemos analizar todas las posibilidades mediante el sistema (6.18): si $y_1 = 0$, se obtiene $x_2 = 80$ y, por tanto, $y_2 = -30$ y $y_3 = -70$ (recuerde que $x_1 = 0$); si $y_2 = 0$, entonces $x_2 = 50$, $y_1 = 30$ y $y_3 = -10$; si $y_3 = 0$, se tiene $x_2 = 45$, $y_1 = 35$ y $y_2 = 5$. Puesto que las variables de holgura tienen que ser no negativas, la única posibilidad es $y_3 = 0$, con lo que se alcanza el punto esquina $(0, 45, 35, 5, 0)$ para el sistema (6.18); es decir, el punto esquina $(0, 45)$ del conjunto factible con los valores de holgura $y_1 = 35$, $y_2 = 5$ y $y_3 = 0$. Ahora bien

$$\begin{aligned} y_1 &= \left(\frac{80}{1} - x_2\right) \\ y_2 &= \left(\frac{50}{1} - x_2\right) \\ y_3 &= \left(\frac{90}{2} - x_2\right) \end{aligned}$$

de donde se observa que la menor razón de $\frac{80}{1}$, $\frac{50}{1}$ y $\frac{90}{2}$, $x_2 = \frac{90}{2} = 45$, satisface la restricción $y_i \geq 0$ para todo i. Para saber cuál es esta razón, utilizando la tabla inicial 6.19, simplemente se calculan los cocientes al dividir cada registro de la última columna entre el registro correspondiente en la misma fila que está en la columna objetivo (exceptuando los registros en la fila objetivo) y se elige de entre éstos el valor mínimo. A la fila en la tabla inicial, que contiene los registros que hacen mínima esta razón, le llamaremos fila pivote. Podemos utilizar la información conocida de que $x_1 = 0$ y $y_3 = 0$ para calcular el punto esquina $(0, 45)$ con el contenido de la tabla inicial 6.19. Empleamos el método de Gauss-Jordan para cambiar la escala del registro contenido en la intersección de la fila pivote y la columna pivote (en este caso 2) a 1 y con este elemento transformar los registros por encima y por debajo de éste en ceros:

$$\begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 80 \\ 1 & 1 & 0 & 1 & 0 & 50 \\ 1 & 2 & 0 & 0 & 1 & 90 \\ -500 & -700 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 80 \\ 1 & 1 & 0 & 1 & 0 & 50 \\ 1/2 & 1 & 0 & 0 & 1/2 & 45 \\ -500 & -700 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 3/2 & 0 & 1 & 0 & -1/2 & 35 \\ 1/2 & 0 & 0 & 1 & -1/2 & 5 \\ 1/2 & 1 & 0 & 0 & 1/2 & 45 \\ -150 & 0 & 0 & 0 & 350 & 31\,500 \end{bmatrix}$$

Dado que $x_1 = y_3 = 0$, tenemos: $x_2 = 45$, $y_1 = 35$ y $y_2 = 5$; es decir, la solución factible $(0, 45, 35, 5, 0)$ del sistema (6.18) con $P = 31\,500$. Esta información la podemos representar en una segunda tabla, análoga a la inicial, que nos proporciona la información del estado del problema en el punto esquina $(0, 45)$:

	x_1	x_2	y_1	y_2	y_3	
y_1	3/2	0	1	0	−1/2	35
y_2	1/2	0	0	1	−1/2	5
x_2	1/2	1	0	0	1/2	45
P	−150	0	0	0	350	31 500

(6.20)

De nuevo, esta segunda tabla se interpreta con el espíritu de la tabla inicial: las variables que están en la primera columna tienen los valores correspondientes en la última columna y las variables de la primera fila que no están en la primera columna toman el valor cero; es decir, $(x_1, x_2, y_1, y_2, y_3) = (0, 45, 35, 5, 0)$ y $P = 31\,500$. A las variables que se encuentran en la primera columna en ambas tablas se les llama variables básicas y a las restantes variables no básicas. En la tabla inicial las variables básicas son y_1, y_2, y_3 y las variables no básicas son x_1 y x_2; mientras que en la tabla para el estado del problema en el punto esquina $(0, 45)$ las variables básicas y no básicas son, respectivamente, y_1, y_2, x_2 y x_1, y_3. Note que otra vez, como en la tabla inicial, los vectores que están por debajo de cada variable básica de la primera fila son los vectores unitarios canónicos $(0, 0, 1, 0)$, $(1, 0, 0, 0)$ y $(0, 1, 0, 0)$ para x_2, y_1 y y_2, respectivamente. De manera completamente análoga, veamos cómo es posible formar la tabla que corresponde al estado del problema en el siguiente punto esquina (que, como sabemos del ejemplo 6.25, es $(10, 40)$) siguiendo una arista para obtener el incremento de P más rápido posible. En la fila objetivo el registro negativo con mayor valor absoluto de la tabla (6.20) es -150; ahora x_1 va a incrementar su valor de cero a un valor mayor a cero y pasará de ser variable no básica a variable básica en la siguiente tabla; así la columna pivote en este estado es la que contiene a x_1. Alguna de las variables de la primera columna de (6.20) pasará de ser variable básica a variable no básica. Para saber cuál es, calculamos los cocientes al dividir los registros de la última columna de (6.20) entre los correspondientes registros de la columna pivote (excepto los de la fila objetivo):

$$\frac{35}{\frac{3}{2}} = \frac{70}{3},$$
$$\frac{5}{\frac{1}{2}} = 10,$$
$$\frac{45}{\frac{1}{2}} = 90;$$

la menor razón es $\frac{5}{1/2}$ que corresponde a la variable y_2. Se dice que y_2 es **variable básica saliente** y x_1 es **variable básica entrante** para establecer la siguiente tabla. Como x_1 será variable básica, los elementos que están por debajo de ella deben corresponder a un vector canónico unitario; para hacer esto utilizamos el elemento que se encuentra en la intersección de la fila pivote y la columna pivote para, con el método de Gauss-Jordan, cambiar la escala de este elemento a 1 y transformar los registros por encima y debajo de éste en ceros:

	x_1	x_2	y_1	y_2	y_3	
y_1	$3/2$	0	1	0	$-1/2$	35
y_2	$1/2$	0	0	1	$-1/2$	5
x_2	$1/2$	1	0	0	$1/2$	45
P	-150	0	0	0	350	31 500

$\sim$

	x_1	x_2	y_1	y_2	y_3	
y_1	0	0	1	-3	1	20
x_1	1	0	0	2	-1	10
x_2	0	1	0	-1	1	40
P	0	0	0	300	200	33 000

De la última tabla obtenemos la solución factible $(10, 40, 20, 0, 0)$ y $P = 33\,000$, es decir, $x_1 = 10$, $x_2 = 40$ para una ganancia $P = 33\,000$. Hemos encontrado la solución del problema, ¿pero cómo saber por este método que efectivamente ya se alcanzó el máximo y debemos detener el proceso? La respuesta es sencilla, la fila objetivo de la última tabla no contiene ya registros negativos; es ésta la señal que nos indica que hemos llegado al valor óptimo, porque ya no tenemos una dirección que seguir a través de una arista para incrementar la función objetivo. En la figura 6-16 se puede observar claramente la razón de este argumento.

Motivados por la precedente discusión resumimos a continuación el algoritmo para resolver este tipo de problemas.

Método simplex para el problema estándar de programación lineal

Para resolver el problema estándar de programación lineal:

$$\text{maximizar} \quad P = c_1x_1 + c_2x_2 + \cdots + c_nx_n \tag{6.21}$$

$$\text{sujeto a} \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \le b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \le b_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \ldots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \le b_m \end{cases} \tag{6.22}$$
$$x_j \ge 0, b_i \ge 0 \quad \forall i, j$$

o en su forma matricial equivalente (6.15) (cfr. pág. 620), se procede de la siguiente manera:

PLE 1: Se incorporan m variables de **holgura** $y_1, y_2, \ldots, y_m \ge 0$ al sistema de desigualdades (6.22) para transformar éste en el sistema de ecuaciones

$$\begin{array}{ccccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & + & y_1 & & = b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & & + \; y_2 & & = b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots & & \vdots \;\; \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & & & + \; y_m & = b_m \end{array} \tag{6.23}$$

A toda solución $(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$ del sistema (6.23) se le nombra **solución factible**; al conjunto de todas sus soluciones **conjunto factible** del problema estándar de programación lineal y toda solución de este sistema que se obtiene al hacer ceros n de las $n+m$ variables, es un **punto esquina** del conjunto factible para las restricciones (6.22).

PLE 2: Se construye la **tabla inicial**

	x_1	x_2	$\cdots$	x_n	y_1	y_2	$\cdots$	y_m	
y_1	a_{11}	a_{12}	$\cdots$	a_{1n}	1	0	$\cdots$	0	b_1
y_2	a_{21}	a_{21}	$\cdots$	a_{2n}	0	1	$\cdots$	0	b_2
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
y_m	a_{m1}	a_{m2}	$\cdots$	a_{mn}	0	0	$\cdots$	1	b_m
P	$-c_1$	$-c_2$	$\cdots$	$-c_n$	0	0	$\cdots$	0	0

(6.24)

o, en forma matricial

	$\vec{x}^t$	$\vec{y}^t$	
$\vec{y}$	A	I_m	$\vec{b}$
P	$-\vec{c}^t$	$\vec{0}^t_{\mathbb{R}^m}$	0

que está conformada por:

1. Primera fila: las variables del problema $x_1, x_2, \ldots, x_n$ y las **variables de holgura** $y_1, y_2, \ldots, y_m$, por encima de la matriz ampliada $[A \mid I_m]$, donde I_m es la matriz identidad de orden m.
2. Primera columna: las variables de holgura $y_1, y_2, \ldots, y_m$ en la columna adyacente a la izquierda de la matriz $[A \mid I_m]$.
3. Última fila: la variable P para la función de ganancia y los negativos $-c_i$ de los coeficientes de las variables x_i de ésta en (6.21); los registros finales de esta fila son ceros.
4. Última columna: los valores $b_1, b_2, \ldots, b_m$ de las restricciones del problema; el registro final de esta columna es cero.

PLE 3: A partir de la tabla inicial se calculan en forma recurrente tablas que darán información del estado del problema en puntos esquina adyacentes para los que la función objetivo se va incrementando. En cada una de estas tablas emplearemos las siguientes connotaciones e interpretaciones.

1. La primera fila es idéntica a la primera fila de la tabla inicial.
2. A la última fila se le llama **fila objetivo**.
3. Las variables de la primera fila que están por encima de columnas que contienen a uno de los vectores canónicos unitarios de la base de $\mathbb{R}^{m+1}$, $\vec{e}_i = (\underbrace{0, \ldots, 0, \overset{i}{1}, 0, \ldots, 0}_{m+1})$, se llaman **variables básicas** y las restantes variables **no básicas**. (Así, en la tabla inicial, las variables de holgura, y_i, son todas básicas y las variables del problema, x_i, son todas no básicas.)
4. Cada variable en la primera columna (excepto la variable P) es básica y tiene el valor que corresponde al registro que está en la fila a la que pertenece y en la última columna. (En la tabla inicial cada variable de holgura es básica y $y_i = b_i$ para todo i.)
5. Las variables no básicas tienen valores nulos. (En la tabla inicial $x_j = 0$ para todo j.)

6. La variable P, para la función de ganancia, toma el valor del registro que se encuentra en la última columna de la fila objetivo. (Así, en la tabla inicial, $P = 0$.)

PLE 4: Para calcular la tabla siguiente $(\nu + 1)$ a partir de una tabla precedente (ν) se procede así:

1. Una de las variables no básicas de la tabla (ν) será variable básica en la tabla $(\nu+1)$; a esta variable se le llama **variable básica entrante**. Para elegir la variable básica entrante se identifica el registro no negativo de mayor valor absoluto en la fila objetivo de la tabla (ν), la variable que se encuentre en la columna donde está ese registro será la variable básica entrante; dicha columna se llama **columna pivote**.
2. Una de las variables básicas, que están en la primera columna de la tabla (ν), será variable no básica en la tabla $(\nu + 1)$, a ésta se le llama **variable básica saliente**. Para decidir cuál es la variable básica saliente se calculan los cocientes, *definidos y no negativos*, de dividir los registros de la última columna entre los registros de la columna objetivo de la tabla (ν) (únicamente los que corresponden a filas que contienen variables básicas); la fila que corresponda a **la menor razón no negativa** es la que contendrá la variable básica saliente.[9] A esta fila se le llama **fila pivote**.
3. Si todos los registros de la fila pivote son cero o negativos, entonces el problema **no es acotado**; es decir, la función de ganancia **no alcanza el máximo**.
4. La variable básica entrante de la tabla (ν) ocupará el lugar de la variable básica saliente en la primera columna de la tabla $(\nu + 1)$.
5. Mediante el método de Gauss-Jordan el registro contenido en la intersección de la fila pivote y la columna pivote se convierte a 1, mediante un cambio de escala y se transforman en ceros los elementos por encima y por debajo de ese registro en la tabla (ν).
6. La tabla obtenida mediante los pasos precedentes será la tabla $(\nu + 1)$. El punto $\vec{x} = (x_1, x_2, \ldots, x_n)$, que corresponde a los valores que contiene la tabla $(\nu + 1)$, es el **punto esquina** que da la información del problema en ese estado; es decir, en dicho punto esquina.

PLE 5: El proceso termina cuando por primera vez se obtiene una tabla que no contiene registros negativos en la fila objetivo. Los valores que toman P y $\vec{x} = (x_1, x_2, \ldots, x_n)$ en esta tabla serán el valor máximo y un punto esquina donde éste se alcanza, respectivamente.

❍ **Nota 6.2**

1. Para poder tener m variables básicas en la construcción de cada tabla al emplear el método simplex, supondremos que la matriz A en (6.15) tiene rango m.
2. Existen algunos casos, generalmente cuando se presentan restricciones redundantes, en los que en el paso PLE 4, del método simplex, para seleccionar una fila pivote, se obtiene una razón nula; esto es, hay una variable básica que es igual a cero. Se dice entonces que la respectiva solución básica $[\vec{x}\ \vec{y}]^t$ de la primera columna, en la tabla correspondiente, es degenerada. Cuando esto sucede es posible que en lugar de que el método simplex produzca una solución óptima, entre en un ciclo infinito y no se pueda encontrar la solución aun si ésta existe. En este libro no trataremos esos casos.

Ilustremos el algoritmo anterior con el ejemplo de un problema de programación lineal estándar de más de dos variables.

[9] Si existen dos o más razónes mínimas no negativas iguales, se puede elegir cualquiera de las filas correspondientes para determinar la variable básica saliente.

► **Ejemplo 6.26** Una empresa fabrica "cocinetas" en tres modelos: M1, M2 y M3. Los tiempos de fabricación para estos modelos son de tres semanas, dos semanas y una semana, respectivamente. Por cuestiones de espacio únicamente puede almacenar a lo más 28 "cocinetas" en total al año y la empresa labora 48 semanas al año. La ganancia neta por unidad para los modelos M1, M2 y M3 son $5 000, $4 000 y $3 000, respectivamente. La empresa necesita saber cuántas unidades de cada modelo debe fabricar para maximizar la ganancia por la venta de estos artículos.◄

Solución Sean x_1, x_2 y x_3 el número de unidades que la empresa debe fabricar de sendos modelos M1, M2 y M3. Entonces la ganancia (en miles) es $P = 5x_1 + 4x_2 + 3x_3$. Dado que la empresa puede almacenar a lo más 28 "cocinetas", labora 48 semanas en un año, el modelo M1 se fabrica en 3 semanas, el M2 en dos semanas y el M3 se manufactura en una semana, entonces x_1, x_2 y x_3 deben satisfacer las restricciones

$$\begin{aligned} x_1 + x_2 + x_3 &\leq 28 \\ 3x_1 + 2x_2 + x_3 &\leq 48 \end{aligned}$$

Es decir, el problema es:

$$\begin{array}{ll} \text{maximizar} & P = 5x_1 + 4x_2 + 3x_3 \\ \text{sujeto a} & \left\{ \begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & \leq & 28 \\ 3x_1 & + & 2x_2 & + & x_3 & \leq & 48 \end{array} \right. \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

Este es un problema estándar de programación lineal. Se incorporan dos variables de holgura (una para cada restricción) para transformar el sistema de desigualdades en el sistema lineal

$$\begin{array}{rcrcrcrcrcl} x_1 & + & x_2 & + & x_3 & + & y_1 & & & = & 28 \\ 3x_1 & + & 2x_2 & + & x_3 & & & + & y_2 & = & 48 \end{array}$$

Formemos la tabla inicial de acuerdo con (6.24) del método simplex para el problema estándar de programación lineal:

	x_1	x_2	x_3	y_1	y_2	
y_1	1	1	1	1	0	28
y_2	3	2	1	0	1	48
P	−5	−4	−3	0	0	0

El registro negativo con mayor valor absoluto de la fila objetivo es -5, por tanto, x_1 es la variable básica entrante y la columna donde se encuentra es la columna pivote; la menor razón de $28/1 = 28$ y $48/3 = 16$ (los cocientes de los registros de la última columna con la columna pivote) es $48/3$, por lo que y_2 es la variable básica saliente y la fila donde se encuentra es la fila pivote. Con el registro que está en la intersección de la fila pivote y la columna pivote de la tabla inicial (3), empleamos el método de Gauss-Jordan para cambiar su escala a 1 y transformar en ceros todos los registros por encima y debajo de él:

	x_1	x_2	x_3	y_1	y_2	
y_1	1	1	1	1	0	28
y_2	3	2	1	0	1	48
P	−5	−4	−3	0	0	0

∼

	x_1	x_2	x_3	y_1	y_2	
y_1	0	1/3	2/3	1	−1/3	12
x_1	1	2/3	1/3	0	1/3	16
P	0	−2/3	−4/3	0	5/3	80

El registro negativo de mayor valor absoluto en la fila objetivo de la última tabla es $-4/3$, por lo que donde se encuentra se tiene la columna pivote y x_3 es la variable básica entrante. Las razones de los registros de la última columna con los de la columna pivote son $12/(2/3) = 18$, $16/(1/3) = 48$, la menor es $12/(2/3)$; entonces la variable básica saliente es y_1 y la fila donde se encuentra es la fila pivote. Al registro en la intersección de la fila pivote y la columna pivote, $2/3$, le cambiamos la escala a 1 y hacemos ceros los elementos por debajo de éste mediante el método de Gauss-Jordan:

	x_1	x_2	x_3	y_1	y_2				x_1	x_2	x_3	y_1	y_2	
y_1	0	1/3	2/3	1	−1/3	12		x_3	0	1/2	1	3/2	−1/2	18
x_1	1	2/3	1/3	0	1/3	16	∼	x_1	1	1/2	0	−1/2	1/2	10
P	0	−2/3	−4/3	0	5/3	80		P	0	0	0	2	1	104

Puesto que la fila objetivo de la última tabla no tiene registros negativos, el proceso ha terminado con el valor máximo $P = 104$ que se alcanza en $x_1 = 10$, $x_3 = 18$ y $x_2 = 0$. Es decir, la empresa debe fabricar $x_1 = 10$ "cocinetas" del modelo M1, $x_2 = 0$ "cocinetas" del modelo M2 y $x_3 = 18$ "cocinetas" del modelo M3 para obtener una ganancia máxima de \$104 000. ✔

▶ Ejemplo 6.27

$$\text{Maximizar} \quad P = 3x_1 + 2x_2 + x_3$$
$$\text{sujeto a} \quad \begin{cases} x_1 + x_2 \le 2 \\ \quad\; x_2 \le 1 \end{cases}$$
$$x_1, x_2, x_3 \ge 0 \;◀$$

Solución El sistema lineal con las correspondientes variables de holgura es

$$\begin{aligned} x_1 + x_2 + y_1 \quad &= 2 \\ x_2 \quad + y_2 &= 1 \end{aligned}$$

y la tabla inicial

	x_1	x_2	x_3	y_1	y_2	
y_1	1	1	0	1	0	2
y_2	0	1	0	0	1	1
P	−3	−2	−1	0	0	0

El registro negativo con mayor valor absoluto en la fila objetivo es -3, así que la columna pivote es la que contiene a x_1; la menor razón positiva de los registros de la última columna a los de la columna pivote es $2/1$, por tanto, la fila pivote es la que contiene y_1. Entonces, la siguiente tabla es

	x_1	x_2	x_3	y_1	y_2	
x_1	1	1	0	1	0	2
y_2	0	1	0	0	1	1
P	0	1	−1	3	0	6

En esta tabla el registro negativo con mayor valor absoluto en la fila objetivo es -1; así que la columna pivote es la que contiene x_3; sin embargo, todos los registros de la columna pivote son ceros, entonces el problema es no acotado; es decir, P no alcanza valor máximo en el conjunto factible. Este hecho es fácil de comprobar: los puntos $(0, 0, x_3)$, $x_3 \ge 0$, satisfacen evidentemente las restricciones y puesto que

$$\lim_{x_3 \to \infty} P(0, 0, x_3) = \lim_{x_3 \to \infty} x_3 = \infty$$

es evidente que P no alcanza máximo en el conjunto factible. ✔

Minimización por medio de la maximización

Si B es cualquier conjunto de números reales utilizamos las notaciones $\text{máx}(B)$ y $\text{mín}(B)$ para representar el máximo y el mínimo de los elementos de B (si es que B los tiene), respectivamente, y si A es un conjunto de números reales, $-A$ denota el conjunto formado por los negativos de los elementos de A. Si uno de $\text{máx}(-A)$ o $\text{mín}(A)$ existe, entonces el otro también y, además,

$$\text{mín}(A) = -\text{máx}(-A) \tag{6.25}$$

En efecto, supongamos que $\text{mín}(A)$ existe, entonces $\text{mín}(A) \in A$ y

$$\text{mín}(A) \leq a \quad \forall\, a \in A,$$

por tanto,

$$-\text{mín}(A) \geq -a \quad \forall\, a \in A,$$

luego,

$$-\text{mín}(A) = \text{máx}(-A);$$

si $\text{máx}(-A)$ existe, entonces $\text{máx}(-A) \in -A$ y

$$\text{máx}(-A) \geq -a \quad \forall\, a \in A,$$

por lo que

$$-\text{máx}(-A) \leq a \quad \forall\, a \in A,$$

y, por tanto,

$$-\text{máx}(-A) = \text{mín}(A).$$

Podemos usar (6.25) para resolver problemas de minimización con la misma técnica de maximización; calculando el máximo de $-C$ y multiplicando este resultado por -1.

▶ **Ejemplo 6.28**

$$\begin{array}{ll} \text{Minimizar} & C = -x_1 + 8x_2 + 4x_3 \\ \text{sujeto a} & \left\{ \begin{array}{rcrcrcr} x_1 & + & x_2 & - & x_3 & \leq & 5 \\ 2x_1 & + & x_2 & + & 3x_3 & \leq & 16 \\ 3x_1 & + & 2x_2 & + & x_3 & \leq & 25 \end{array} \right. \\ & x_1, x_2, x_3 \geq 0 \end{array} \tag{6.26}$$

Solución El problema se puede resolver maximizando la función $P = -C = x_1 - 8x_2 - 4x_3$ sujeta a las mismas restricciones. La tabla inicial es entonces

	x_1	x_2	x_3	y_1	y_2	y_3	
y_1	1	1	-1	1	0	0	5
y_2	2	1	3	0	1	0	16
y_3	3	2	1	0	0	1	25
P	-1	8	4	0	0	0	0

El registro negativo con mayor valor absoluto en la fila objetivo es -1 y la menor razón de la última columna en la columna objetivo es $5/1$; así que sale la variable y_1 de la primera columna y entra la variable x_1; hacemos un pivote con el método de Gauss-Jordan para convertir a ceros los registros por debajo del elemento que se encuentra en la intersección de la fila pivote y la columna pivote:

	x_1	x_2	x_3	y_1	y_2	y_3	
y_1	1	1	-1	1	0	0	5
y_2	2	1	3	0	1	0	16
y_3	3	2	1	0	0	1	25
P	-1	8	4	0	0	0	0

$\sim$

	x_1	x_2	x_3	y_1	y_2	y_3	
x_1	1	1	-1	1	0	0	5
y_2	0	-1	5	-2	1	0	6
y_3	0	-1	4	-3	0	1	10
P	0	9	3	1	0	0	5

Dado que ya no hay registros negativos en la fila objetivo, el proceso ha terminado. Así que el máximo de $P = -C$ es 5 y se alcanza en $(5, 0, 0)$. Por tanto, el mínimo de $C = -x_1 + 8x_2 + 4x_3$ sujeto a las restricciones (6.26) es $C = -5$ y se alcanza en $x_1 = 5$, $x_2 = 0$ y $x_3 = 0$. ✔

6.4.3 Restricciones generales y método simplex de dos fases

Antes de atacar el problema de programación lineal con restricciones generales de los tipos (6.13) (cfr. pág. 620), observemos lo siguiente:

S1. Todo problema de minimización de una función de costo C se puede resolver maximizando la función de ganancia $P = -C$.

S2. Toda restricción del tipo

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \geq b_i$$

se puede transformar en la restricción equivalente

$$-a_{i1}x_1 - a_{i2}x_2 - \cdots - a_{in}x_n \leq -b_i.$$

S3. Cualquier restricción de la forma

$$a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = b_k$$

equivale al par de restricciones

$$\begin{aligned} a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n &\leq b_k, \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n &\geq b_k. \end{aligned}$$

Entonces, para solucionar cualquier problema de programación lineal con restricciones generales (6.13), basta saber cómo se resuelve el problema de programación lineal

$$\begin{array}{ll} \text{maximizar} & P = \vec{c}^{\,t}\vec{x} \\ \text{sujeto a} & A\vec{x} \leq \vec{b} \\ & \vec{x} \geq \vec{0}_{\mathbb{R}^n} \end{array} \tag{6.27}$$

donde $\vec{b} = [\; b_1 \quad b_2 \quad \cdots \quad b_m \;]^t$ y los $b_i \in \mathbb{R}$; es decir, los números b_i pueden ser negativos.

Para resolver el problema (6.27) utilizaremos una técnica que consta de dos etapas o fases. En la primera, a partir de la información en el planteamiento original, se construye una tabla que contiene el estado del problema en un punto esquina del conjunto factible con los elementos de la última columna no negativos; en la segunda fase simplemente se aplica el método simplex para el problema estándar tomando como tabla inicial aquella construida en la primera fase. Vamos a ilustrar las ideas centrales de este método con el siguiente ejemplo.

► **Ejemplo 6.29** Resolver el problema de programación lineal

$$\begin{array}{ll} \text{maximizar} & P = 3x_1 + 5x_2 \\ \text{sujeto a} & \left\{\begin{array}{rcrcr} -x_1 & + & 5x_2 & \leq & 29 \\ x_1 & + & x_2 & \geq & 7 \\ 3x_1 & + & 2x_2 & \leq & 32 \\ x_1 & + & 3x_2 & \geq & 13 \end{array}\right. \\ & x_1, x_2 \geq 0 \end{array} \tag{6.28}$$
◄

Solución **Etapa I:** Construcción de una tabla inicial:

1. Transformemos la segunda y cuarta restricciones en (6.28), que son del tipo $\geq$, en restricciones de la forma $\leq$ para obtener el problema equivalente

$$\begin{array}{ll} \text{maximizar} & P = 3x_1 + 5x_2 \\ \text{sujeto a} & \left\{\begin{array}{rcrcr} -x_1 & + & 5x_2 & \leq & 29 \\ -x_1 & - & x_2 & \leq & -7 \\ 3x_1 & + & 2x_2 & \leq & 32 \\ -x_1 & - & 3x_2 & \leq & -13 \end{array}\right. \\ & x_1, x_2 \geq 0 \end{array} \tag{6.29}$$

2. Incorporemos, de manera análoga a como se hace en el problema estándar, variables de holgura $y_1, y_2, y_3, y_4 \geq 0$ para transformar las restricciones (6.29) en el sistema de ecuaciones

$$\begin{array}{rcrcccccccccr} -x_1 & + & 5x_2 & + & y_1 & & & & & & & = & 29 \\ -x_1 & - & x_2 & & & + & y_2 & & & & & = & -7 \\ 3x_1 & + & 2x_2 & & & & & + & y_3 & & & = & 32 \\ -x_1 & - & 3x_2 & & & & & & & + & y_4 & = & -13 \end{array} \tag{6.30}$$

3. Para aplicar el método simplex como se hace en el caso estándar necesitamos:

   (a) que los términos independientes de (6.30) en el lado derecho de los signos de igualdad sean positivos,

   (b) una tabla inicial con el estado del problema en un punto esquina del conjunto factible.

Para alcanzar este objetivo, construyamos la siguiente matriz, que contiene la matriz ampliada del sistema (6.30) y en su última fila, a la fila objetivo (sin la etiqueta P) de este problema tal como se forma en el caso estándar:

$$\begin{bmatrix} -1 & 5 & 1 & 0 & 0 & 0 & 29 \\ -1 & -1 & 0 & 1 & 0 & 0 & -7 \\ 3 & 2 & 0 & 0 & 1 & 0 & 32 \\ -1 & -3 & 0 & 0 & 0 & 1 & -13 \\ -3 & -5 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Con ella vamos a seleccionar el registro negativo que se encuentre más abajo en su última columna y elegimos, en la fila que lo contiene, un elemento negativo para transformarlo a 1 mediante cambio de escala y con él convertir en ceros los elementos que están arriba y abajo mediante el método de Gauss-Jordan (en este caso se eligió a -1):

$$\begin{bmatrix} -1 & 5 & 1 & 0 & 0 & 0 & 29 \\ -1 & -1 & 0 & 1 & 0 & 0 & -7 \\ 3 & 2 & 0 & 0 & 1 & 0 & 32 \\ -1 & -3 & 0 & 0 & 0 & 1 & -13 \\ -3 & -5 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} -1 & 5 & 1 & 0 & 0 & 0 & 29 \\ -1 & -1 & 0 & 1 & 0 & 0 & -7 \\ 3 & 2 & 0 & 0 & 1 & 0 & 32 \\ 1 & 3 & 0 & 0 & 0 & -1 & 13 \\ -3 & -5 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 0 & 8 & 1 & 0 & 0 & -1 & 42 \\ 0 & 2 & 0 & 1 & 0 & -1 & 6 \\ 0 & -7 & 0 & 0 & 1 & 3 & -7 \\ 1 & 3 & 0 & 0 & 0 & -1 & 13 \\ 0 & 4 & 0 & 0 & 0 & -3 & 39 \end{bmatrix}$$

Repetimos lo anterior ahora con el registro -7 de la última columna en la matriz precedente; pero, para asegurarnos que los registros en la última columna por debajo de -7 continúen positivos, calculamos las razones no negativas y definidas de los registros de la última columna con los registros de la segunda columna que están por debajo de -7, a partir del propio -7 (sin incluir en este paso la última fila) y elegimos la menor; el registro en la fila de la segunda columna que corresponde a la razón mínima se utiliza para cambiar su escala a 1 y transformar los elementos por debajo y encima de él en ceros, con el método de Gauss-Jordan. De $-7/(-7)$, $13/3$ la razón mínima es $-7/(-7)$, entonces:

$$\begin{bmatrix} 0 & 8 & 1 & 0 & 0 & -1 & 42 \\ 0 & 2 & 0 & 1 & 0 & -1 & 6 \\ 0 & -7 & 0 & 0 & 1 & 3 & -7 \\ 1 & 3 & 0 & 0 & 0 & -1 & 13 \\ 0 & 4 & 0 & 0 & 0 & -3 & 39 \end{bmatrix} \sim \begin{bmatrix} 0 & 8 & 1 & 0 & 0 & -1 & 42 \\ 0 & 2 & 0 & 1 & 0 & -1 & 6 \\ 0 & 1 & 0 & 0 & -1/7 & -3/7 & 1 \\ 1 & 3 & 0 & 0 & 0 & -1 & 13 \\ 0 & 4 & 0 & 0 & 0 & -3 & 39 \end{bmatrix}$$

$$\sim \begin{bmatrix} 0 & 0 & 1 & 0 & 8/7 & 17/7 & 34 \\ 0 & 0 & 0 & 1 & 2/7 & -1/7 & 4 \\ 0 & 1 & 0 & 0 & -1/7 & -3/7 & 1 \\ 1 & 0 & 0 & 0 & 3/7 & 2/7 & 10 \\ 0 & 0 & 0 & 0 & 4/7 & -9/7 & 35 \end{bmatrix} \qquad (6.31)$$

Con lo que obtenemos el punto esquina $(x_1, x_2) = (10, 1)$ para una ganancia $P = 35$; es decir, la solución factible $(x_1, x_2, y_1, y_2, y_3) = (10, 1, 34, 4, 0)$ del sistema (6.30). Con la matriz (6.31) ahora podemos construir la tabla inicial para el estado en el punto esquina (10, 1); es decir, en la solución factible $(x_1, x_2, y_1, y_2, y_3) = (10, 1, 34, 4, 0)$ del sistema (6.30):

	x_1	x_2	y_1	y_2	y_3	y_4	
y_1	0	0	1	0	8/7	17/7	34
y_2	0	0	0	1	2/7	−1/7	4
x_2	0	1	0	0	−1/7	−3/7	1
x_1	1	0	0	0	3/7	2/7	10
P	0	0	0	0	4/7	−9/7	35

(6.32)

Etapa II: Apliquemos ahora directamente el método simplex del caso estándar tomando como tabla inicial la tabla (6.32):

	x_1	x_2	y_1	y_2	y_3	y_4	
y_1	0	0	1	0	8/7	17/7	34
y_2	0	0	0	1	2/7	−1/7	4
x_2	0	1	0	0	−1/7	−3/7	1
x_1	1	0	0	0	3/7	2/7	10
P	0	0	0	0	4/7	−9/7	35

∼

	x_1	x_2	y_1	y_2	y_3	y_4	
y_1	0	0	7/17	0	8/17	1	14
y_2	0	0	0	1	2/7	−1/7	4
x_2	0	1	0	0	−1/7	−3/7	1
x_1	1	0	0	0	3/7	2/7	10
P	0	0	0	0	4/7	−9/7	35

∼

	x_1	x_2	y_1	y_2	y_3	y_4	
y_4	0	0	7/17	0	8/17	1	14
y_2	0	0	1/17	1	6/17	0	6
x_2	0	1	3/17	0	1/17	0	7
x_1	1	0	−2/17	0	5/17	0	6
P	0	0	9/17	0	20/17	0	53

De donde se obtiene la solución $(x_1, x_2) = (6, 7)$ para un máximo $P = 53$. ✔

A continuación resumimos el algoritmo que motivamos con el ejemplo precedente.

Método simplex de dos fases

Para resolver el problema[10]

$$\begin{array}{ll} \text{maximizar} & P = \vec{c}^{\,t}\vec{x} \\ \text{sujeto a } \mu \text{ restricciones del tipo} & \{\sum_{j=1}^{n} \alpha_{ij}x_j \left\{ \begin{array}{ll} \leq & \beta_i \quad o \\ = & \beta_i \quad o \\ \geq & \beta_i \end{array} \right. \\ & x_j \geq 0, b_j \in \mathbb{R}, \forall i, j \end{array} \tag{6.33}$$

se procede de la siguiente manera:

Fase I: Construcción de una tabla inicial.

1. *(a)* Toda restricción del tipo $=$ en (6.33) se transforma en un par de restricciones equivalentes de los tipos $\leq$ y $\geq$ (cfr. S3, pág. 631).
   (b) Todas las restricciones que resulten con el símbolo $\geq$ en el paso *(a)* y todas las restricciones de la forma $\geq$ en (6.33) se transforman en restricciones del tipo $\leq$ (cfr. S2, pág. 631) para obtener un sistema de restricciones equivalente de la forma

$$\sum_{j=1}^{n} a_{ij}x_j \leq b_i, \quad i = 1, \ldots, m \tag{6.34}$$

2. Para cada $i = 1, \ldots, m$ se incorpora a la restricción (6.34) una variable de holgura $y_i \geq 0$ para transformar esta restricción en la ecuación

[10] Recuerde que para hallar el mínimo de una función de costo C, se calcula el máximo de $-C$ y el resultado se multiplica por -1 (cfr. pág. 630 y S1 de la pág. 631).

$$\sum_{j=1}^{n} a_{ij}x_j + y_i = b_i \tag{6.35}$$

3. Se forma la matriz

$$\begin{bmatrix} A & I_m & \vec{b} \\ -\vec{c}^{\,t} & \vec{0}^{\,t}_{\mathbb{R}^m} & 0 \end{bmatrix} \tag{6.36}$$

donde $\begin{bmatrix} A & I_m & \vec{b} \end{bmatrix}$ es la matriz ampliada de coeficientes del sistema (6.35), esto es, $A = [a_{ij}]$, I_m es la matriz identidad de orden m y $\vec{b} = \begin{bmatrix} b_1 & b_2 & \cdots & b_m \end{bmatrix}^t$; $\vec{c} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}^t$ y $\vec{0}_{\mathbb{R}^m} = [\underbrace{0 \quad 0 \quad \cdots \quad 0}_{m}]^t$.

4. Sea b_k el último elemento negativo de la columna que contiene a $\vec{b}$ de la matriz (6.36). Se elige un registro negativo en la fila donde está b_k, digamos a_{kj} y se calculan las razones no negativas y definidas de los registros en la columna que contiene a $\vec{b}$ con los registros de la columna que contiene a_{kj}, de b_k hacia abajo; esto es, las razones b_l/a_{lj}, $l = k, \ldots, m$, que resulten no negativas y estén definidas. Al registro a_{lj} que corresponde a la mínima razón se le cambia la escala a 1 y con el método de Gauss-Jordan se transforman en ceros los elementos por encima y por debajo de él obteniendo una matriz equivalente por filas a la matriz (6.36).
5. Se aplica el proceso anterior de manera iterativa hasta obtener, por primera vez, que los primeros m elementos en la última columna sean no negativos, resultando una matriz equivalente a todas las precedentes de la forma

$$\begin{bmatrix} A_\nu & \vec{b}_\nu \\ \vec{c}^{\,t}_\nu & \pi \end{bmatrix} \tag{6.37}$$

con $A_\nu \in \mathfrak{M}_{m\times n}$, $\vec{b}_\nu \geq \vec{0}_{\mathbb{R}^m}$ y $\pi \in \mathbb{R}$.

6. La solución $(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$ del sistema $[A_\nu \,|\, \vec{b}_\nu]$ es una solución factible del sistema equivalente (6.35); donde las variables que corresponden a columnas que no son vectores unitarios canónicos (que no son variables básicas) tienen valor cero y las restantes (las básicas) toman el valor que se encuentra en la última columna de (6.37) en la fila que tiene el único uno del vector canónico por debajo de la variable básica. Por tanto $(x_1, x_2, \ldots, x_n)$, es un punto esquina del conjunto factible y $P = \pi$ en este punto esquina.
7. Con la información del inciso anterior y con la matriz (6.37) se construye la tabla inicial para el estado del problema en el punto esquina $(x_1, x_2, \ldots, x_n)$ hallado en el paso anterior.

Fase II: Se aplica el método simplex para el caso estándar a la tabla inicial encontrada en la fase I.

► **Ejemplo 6.30** Resolver el problema

$$\begin{array}{ll} \text{maximizar} & P = 3x_1 + 2x_2 - x_3 \\ \text{sujeto a} & \left\{ \begin{array}{rcrcrcr} x_1 & + & 3x_2 & + & x_3 & \leq & 9 \\ 2x_1 & + & 3x_2 & - & x_3 & \geq & 2 \\ 3x_1 & - & 2x_2 & + & x_3 & \geq & 5 \end{array} \right. \\ & x_1, x_2, x_3 \geq 0 \end{array} \tag{6.38}$$ ◄

Solución **Fase I:** Las restricciones $2x_1 + 3x_2 - x_3 \geq 2$ y $3x_1 - 2x_2 + x_3 \geq 5$ en (6.38) se transforman en $-2x_1 - 3x_2 + x_3 \leq -2$ y $-3x_1 + 2x_2 - x_3 \leq -5$, respectivamente. Al incorporar las variables de holgura se obtiene el sistema

$$\begin{array}{rcrcrcrcrcrcr} x_1 & + & 3x_2 & + & x_3 & + & y_1 & & & & & = & 9 \\ -2x_1 & - & 3x_2 & + & x_3 & & & + & y_2 & & & = & -2 \\ -3x_1 & + & 2x_2 & - & x_3 & & & & & + & y_3 & = & -5 \end{array}$$

Entonces, la matriz que corresponde, en este caso, con el inciso 3 de la fase I del método simplex es

$$\begin{bmatrix} 1 & 3 & 1 & 1 & 0 & 0 & 9 \\ -2 & -3 & 1 & 0 & 1 & 0 & -2 \\ -3 & 2 & -1 & 0 & 0 & 1 & -5 \\ -3 & -2 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

El último registro negativo de la última columna en la matriz anterior es -5; elijamos el registro -1 que se encuentra en la misma fila (la única razón no negativa para este caso es $-5/(-1) = 5$). Entonces, al cambiar la escala de este registro a 1 y aplicar el método de Gauss-Jordan se obtiene

$$\begin{bmatrix} 1 & 3 & 1 & 1 & 0 & 0 & 9 \\ -2 & -3 & 1 & 0 & 1 & 0 & -2 \\ -3 & 2 & -1 & 0 & 0 & 1 & -5 \\ -3 & -2 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 1 & 1 & 0 & 0 & 9 \\ -2 & -3 & 1 & 0 & 1 & 0 & -2 \\ 3 & -2 & 1 & 0 & 0 & -1 & 5 \\ -3 & -2 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$
$$\sim \begin{bmatrix} -2 & 5 & 0 & 1 & 0 & 1 & 4 \\ -5 & -1 & 0 & 0 & 1 & 1 & -7 \\ 3 & -2 & 1 & 0 & 0 & -1 & 5 \\ -6 & 0 & 0 & 0 & 0 & 1 & -5 \end{bmatrix}.$$

Repetimos el paso anterior ahora con el registro -7 de la última columna en la matriz anterior. Tomando a -1 para registro electo en la fila donde se encuentra -7. Como no hay razones positivas por debajo de este registro (de -1), cambiamos la escala a 1 del elemento -1 y transformamos los registros por encima y por debajo de él a ceros mediante el método de Gauss-Jordan:

$$\begin{bmatrix} -2 & 5 & 0 & 1 & 0 & 1 & 4 \\ -5 & -1 & 0 & 0 & 1 & 1 & -7 \\ 3 & -2 & 1 & 0 & 0 & -1 & 5 \\ -6 & 0 & 0 & 0 & 0 & 1 & -5 \end{bmatrix} \sim \begin{bmatrix} -2 & 5 & 0 & 1 & 0 & 1 & 4 \\ 5 & 1 & 0 & 0 & -1 & -1 & 7 \\ 3 & -2 & 1 & 0 & 0 & -1 & 5 \\ -6 & 0 & 0 & 0 & 0 & 1 & -5 \end{bmatrix}$$
$$\sim \begin{bmatrix} -27 & 0 & 0 & 1 & 5 & 6 & -31 \\ 5 & 1 & 0 & 0 & -1 & -1 & 7 \\ 13 & 0 & 1 & 0 & -2 & -3 & 19 \\ -6 & 0 & 0 & 0 & 0 & 1 & -5 \end{bmatrix}.$$

Efectuamos el mismo proceso ahora con el registro -31 de la última columna en la matriz precedente: la menor razón de $-31/(-27)$, $7/5$ y $19/13$ es $7/5$, entonces

$$\begin{bmatrix} -27 & 0 & 0 & 1 & 5 & 6 & -31 \\ 5 & 1 & 0 & 0 & -1 & -1 & 7 \\ 13 & 0 & 1 & 0 & -2 & -3 & 19 \\ -6 & 0 & 0 & 0 & 0 & 1 & -5 \end{bmatrix} \sim \begin{bmatrix} -27 & 0 & 0 & 1 & 5 & 6 & -31 \\ 1 & 1/5 & 0 & 0 & -1/5 & -1/5 & 7/5 \\ 13 & 0 & 1 & 0 & -2 & -3 & 19 \\ -6 & 0 & 0 & 0 & 0 & 1 & -5 \end{bmatrix}$$
$$\sim \begin{bmatrix} 0 & 27/5 & 0 & 1 & -2/5 & 3/5 & 34/5 \\ 1 & 1/5 & 0 & 0 & -1/5 & -1/5 & 7/5 \\ 0 & -13/5 & 1 & 0 & 3/5 & -2/5 & 4/5 \\ 0 & 6/5 & 0 & 0 & -6/5 & -1/5 & 17/5 \end{bmatrix}.$$

La solución factible es, por tanto $(x_1, x_2, x_3, y_1, y_2, y_3) = (7/5, 0, 4/5, 34/5, 0, 0)$ y la tabla inicial

	x_1	x_2	x_3	y_1	y_2	y_3	
y_1	0	27/5	0	1	−2/5	3/5	34/5
x_1	1	1/5	0	0	−1/5	−1/5	7/5
x_3	0	−13/5	1	0	3/5	−2/5	4/5
P	0	6/5	0	0	−6/5	−1/5	17/5

Fase II: Aplicamos el método simplex para el problema estándar a partir de la tabla inicial anterior:

	x_1	x_2	x_3	y_1	y_2	y_3	
y_1	0	27/5	0	1	−2/5	3/5	34/5
x_1	1	1/5	0	0	−1/5	−1/5	7/5
x_3	0	−13/5	1	0	3/5	−2/5	4/5
P	0	6/5	0	0	−6/5	−1/5	17/5

$\sim$

	x_1	x_2	x_3	y_1	y_2	y_3	
y_1	0	11/3	2/3	1	0	1/3	22/3
x_1	1	−2/3	1/3	0	0	−1/3	5/3
y_2	0	−13/3	5/3	0	1	−2/3	4/3
P	0	−4	2	0	0	−1	5

$\sim$

	x_1	x_2	x_3	y_1	y_2	y_3	
x_2	0	1	2/11	3/11	0	1/11	2
x_1	1	0	5/11	2/11	0	−3/11	3
y_2	0	0	27/11	13/11	1	−3/11	10
P	0	0	30/11	12/11	0	−7/11	13

$\sim$

	x_1	x_2	x_3	y_1	y_2	y_3	
y_3	0	11	2	3	0	1	22
x_1	1	3	1	1	0	0	9
y_2	0	3	3	2	1	0	16
P	0	7	4	3	0	0	27

de donde se desprende que el máximo es $P = 27$ y se alcanza en $x_1 = 9$, $x_2 = 0$ y $x_3 = 0$. ✔

► **Ejemplo 6.31** Resolver el problema

$$\begin{array}{ll} \text{maximizar} & P = 3x_1 + 2x_2 \\ \text{sujeto a} & \left\{ \begin{array}{ccccc} 2x_1 & + & x_2 & \leq & 4 \\ x_1 & + & x_2 & = & 3 \end{array} \right. \\ & x_1, x_2 \geq 0 \end{array} \tag{6.39}$$ ◄

Solución **Fase I:** Utilizamos el sistema de restricciones

$$x_1 + x_2 \leq 3$$
$$x_1 + x_2 \geq 3$$

equivalente a la restricción $x_1 + x_2 = 3$ en lugar de ésta. Entonces, el sistema de restricciones (6.39) equivale a

$$\begin{array}{rcrcr} 2x_1 & + & x_2 & \leq & 4 \\ x_1 & + & x_2 & \leq & 3 \\ -x_1 & - & x_2 & \leq & -3 \end{array}$$

Por tanto, la matriz con la que se tiene que trabajar para obtener una tabla inicial es

$$\begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 4 \\ 1 & 1 & 0 & 1 & 0 & 3 \\ -1 & -1 & 0 & 0 & 1 & -3 \\ -3 & -2 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Luego, siguiendo el procedimiento en su primera fase, se tiene

$$\begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 4 \\ 1 & 1 & 0 & 1 & 0 & 3 \\ -1 & -1 & 0 & 0 & 1 & -3 \\ -3 & -2 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & -1 & 1 & 0 & 2 & -2 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & -1 & 3 \\ 0 & 1 & 0 & 0 & -3 & 9 \end{bmatrix}$$

$$\sim \begin{bmatrix} 0 & 1 & -1 & 0 & -2 & 2 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & -1 & 7 \end{bmatrix}$$

La tabla inicial es

	x_1	x_2	y_1	y_2	y_3	
x_2	0	1	−1	0	−2	2
y_2	0	0	0	1	1	0
x_1	1	0	1	0	1	1
P	0	0	1	0	−1	7

Fase II: Apliquemos la fase II a la matriz inicial encontrada en el paso anterior:

	x_1	x_2	y_1	y_2	y_3	
x_2	0	1	−1	0	−2	2
y_2	0	0	0	1	1	0
x_1	1	0	1	0	1	1
P	0	0	1	0	−1	7

$\sim$

	x_1	x_2	y_1	y_2	y_3	
x_2	0	1	−1	2	0	2
y_3	0	0	0	1	1	0
x_1	1	0	1	−1	0	1
P	0	0	1	1	0	7

con lo que se obtiene el valor máximo $P = 7$ que se alcanza en $x_1 = 1$ y $x_2 = 2$. ✔

► **Ejemplo 6.32** Una empresa tiene dos plantas, P_1 y P_2, donde fabrica motores que envía a dos plantas de ensamblaje de automóviles, E_1 y E_2. La siguiente tabla contiene la información de los costos netos (en dólares) por transportación de cada unidad que se envía de la planta de producción P_i a la planta de ensamblaje E_j:

	E_1	E_2
P_1	250	300
P_2	400	350

.

P_1 puede fabricar a lo más 700 motores por semana y P_2 puede fabricar a lo sumo 500 motores por semana. La planta E_1 necesita que al menos se le envíen 600 motores a la semana y E_2 requiere de al menos 400 motores por semana. Encontrar la cantidad que cada planta de producción debe mandar por semana a cada planta de ensamblaje, para que la empresa satisfaga la demanda necesaria a un costo mínimo.◄

Solución Sean

x_1 : cantidad de motores por semana que envía P_1 a E_1,

x_2 : cantidad de motores por semana que envía P_1 a E_2,

x_3 : cantidad de motores por semana que envía P_2 a E_1,

x_4 : cantidad de motores por semana que envía P_2 a E_2.

Entonces, de la tabla anterior, el costo por transporte es

$$C = 250x_1 + 300x_2 + 400x_3 + 350x_4.$$

Dado que la planta P_1 produce a lo más 700 motores y la planta P_2 fabrica a lo sumo 500 motores por semana, se debe tener

$$\begin{aligned} x_1 + x_2 &\leq 700 \quad \text{y} \\ x_3 + x_4 &\leq 500. \end{aligned}$$

Como E_1 y E_2 deben recibir al menos 600 y 400 motores, respectivamente, a la semana, se debe cumplir

$$\begin{aligned} x_1 + x_3 &\geq 600 \text{ y} \\ x_2 + x_4 &\geq 400. \end{aligned}$$

Entonces, el problema consiste en

$$\begin{array}{ll} \text{minimizar} & C = 250x_1 + 300x_2 + 400x_3 + 350x_4 \\ \text{sujeto a} & \left\{ \begin{array}{lcll} x_1 & + & x_2 & \leq 700 \\ x_3 & + & x_4 & \leq 500 \\ x_1 & + & x_3 & \geq 600 \\ x_2 & + & x_4 & \geq 400 \end{array} \right. \\ & x_1, x_2, x_3, x_4 \geq 0 \end{array} \tag{6.40}$$

Fase I: Transformemos las desigualdades del tipo $\geq$ de (6.40) en el sentido $\leq$ para obtener las restricciones equivalentes

$$\begin{array}{rcrcr} x_1 & + & x_2 & \leq & 700 \\ x_3 & + & x_4 & \leq & 500 \\ -x_1 & - & x_3 & \leq & -600 \\ -x_2 & - & x_4 & \leq & -400 \end{array}$$

Al incorporar las variables de holgura y la fila objetivo[11] obtenemos la matriz (6.36), para este caso,

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 700 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 500 \\ -1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & -600 \\ 0 & -1 & 0 & -1 & 0 & 0 & 0 & 1 & -400 \\ 250 & 300 & 400 & 350 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

[11] Note que es un problema de minimización y, por tanto, $\text{mín}(C) = -\text{máx}(-C)$.

Ahora continuamos con el proceso de la fase I para obtener una tabla inicial para este problema:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 700 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 500 \\ -1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & -600 \\ 0 & -1 & 0 & -1 & 0 & 0 & 0 & 1 & -400 \\ 250 & 300 & 400 & 350 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & -1 & 1 & 0 & 0 & 1 & 300 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 500 \\ -1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & -600 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & -1 & 400 \\ 250 & 0 & 400 & 50 & 0 & 0 & 0 & 300 & -120\,000 \end{bmatrix}$$

$$\sim \begin{bmatrix} 0 & 0 & -1 & -1 & 1 & 0 & 1 & 1 & -300 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 500 \\ 1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 600 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & -1 & 400 \\ 0 & 0 & 150 & 50 & 0 & 0 & 250 & 300 & -270\,000 \end{bmatrix}$$

$$\sim \begin{bmatrix} 0 & 0 & 1 & 1 & -1 & 0 & -1 & -1 & 300 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 200 \\ 1 & 0 & 0 & -1 & 1 & 0 & 0 & 1 & 300 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & -1 & 400 \\ 0 & 0 & 0 & -100 & 150 & 0 & 400 & 450 & -315\,000 \end{bmatrix} \tag{6.41}$$

Puesto que los primeros cuatro registros en la última columna de la matriz (6.41) son no negativos, la tabla inicial para este problema es

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	
x_3	0	0	1	1	−1	0	−1	−1	300
y_2	0	0	0	0	1	1	1	1	200
x_1	1	0	0	−1	1	0	0	1	300
x_2	0	1	0	1	0	0	0	−1	400
P	0	0	0	−100	150	0	400	450	−315 000

(6.42)

Fase II: Continuamos con la fase II para este problema tomando como tabla inicial (6.42),

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	
x_3	0	0	1	1	−1	0	−1	−1	300
y_2	0	0	0	0	1	1	1	1	200
x_1	1	0	0	−1	1	0	0	1	300
x_2	0	1	0	1	0	0	0	−1	400
P	0	0	0	−100	150	0	400	450	−315 000

$\sim$

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	
x_4	0	0	1	1	−1	0	−1	−1	300
y_2	0	0	0	0	1	1	1	1	200
x_1	1	0	1	0	0	0	−1	0	600
x_2	0	1	−1	0	1	0	1	0	100
P	0	0	100	0	50	0	300	350	−285 000

Entonces, $x_1 = 600$, $x_2 = 100$, $x_3 = 0$, $x_4 = 300$ y $P = -285\,000$. Es decir, P_1 debe enviar 600 motores a E_1 y 100 motores a E_2; mientras que P_2 debe enviar 300 motores a E_2 y ningún motor a E_1, con un costo mínimo total de \$285 000. ✔

6.4.4 Dualidad

Todo problema estándar de programación lineal tiene asociado un problema dual de programación lineal, de tal manera que si se conoce la solución de uno de ellos, también se conoce la solución del otro. Para motivar este concepto usaremos el siguiente ejemplo.

► **Ejemplo 6.33** Se requiere satisfacer las necesidades alimenticias de vitaminas A, B y C para cierto animal de granja. Para ello existen dos tipos de alimentos, F_1 y F_2, que se les da mezclados a comer a estos animales. Las unidades por gramo de cada vitamina en cada alimento, los costos de cada alimento por gramo y los requerimientos mínimos diarios de vitaminas están contenidos en la siguiente tabla:

Vitamina	F_1	F_2	Requerimientos
A	1 unidad/g	2 unidades/g	120 unidades
B	2 unidades/g	3 unidades/g	180 unidades
C	3 unidades/g	1 unidad/g	150 unidades
Costos	¢ 10	¢ 12	

(6.43)

Hallar las cantidades en gramos que se tienen que comprar de cada alimento para satisfacer los requerimientos vitamínicos a un costo mínimo.◄

Solución Sean y_1 y y_2, las cantidades (en gramos) que se compran de sendos alimentos F_1 y F_2. Entonces, la función de costo (en centavos) para esta compra es

$$C = 10y_1 + 12y_2.$$

De la tabla precedente se observa que y_1 y y_2 deben satisfacer

$$\begin{aligned} y_1 + 2y_2 &\geq 120, \\ 2y_1 + 3y_2 &\geq 180 \quad \text{y} \\ 3y_1 + y_2 &\geq 150. \end{aligned}$$

Por tanto, el problema a resolver es

$$\begin{array}{ll} \text{minimizar} & C = 10y_1 + 12y_2 \\ \text{sujeto a} & \left\{ \begin{array}{rcrcl} y_1 & + & 2y_2 & \geq & 120 \\ 2y_1 & + & 3y_2 & \geq & 180 \\ 3y_1 & + & y_2 & \geq & 150 \end{array} \right. \\ & y_1, y_2 \geq 0 \end{array} \qquad (6.44)$$

En la figura 6-17 se encuentra el bosquejo del conjunto factible y los puntos esquina para las restricciones (6.44) y, la siguiente tabla, contiene la evaluación de la función de costo en dichos puntos:

Punto esquina	$C(y_1, y_2) = 10y_1 + 12y_2$
$(0, 150)$	1800
$(36, 42)$	864
$(120, 0)$	1200

Por lo que el costo mínimo es $C = 864$ y se alcanza en $(y_1, y_2) = (36, 42)$; es decir, se deben comprar $y_1 = 36$ gramos de alimento F_1 y $y_2 = 42$ gramos de alimento F_2 para un costo mínimo de \$8.64 ✔

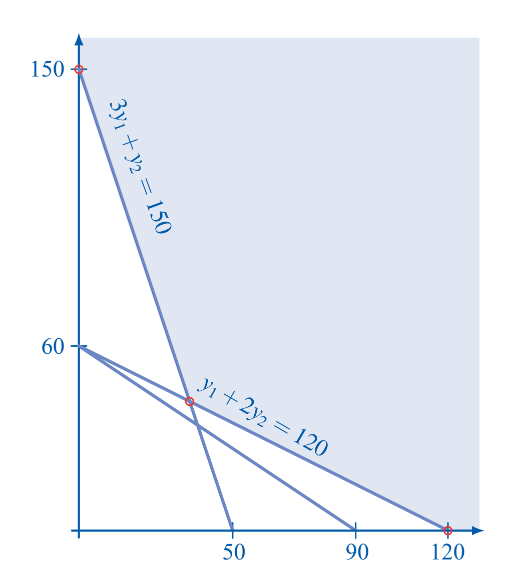


Figura 6-17 •

Pensemos nuevamente en el ejemplo precedente, pero desde el punto de vista de un competidor de diferente giro. Supongamos que también una farmacia veterinaria quiere competir vendiendo estas vitaminas de producción sintética. Sean x_1, x_2 y x_3 los precios (en centavos) por unidad de las vitaminas A, B y C, respectivamente, que desea determinar la farmacia para optimizar su ganancia. Entonces, puesto que se requieren al menos (y con ello es suficiente) 120, 180 y 150 unidades de sendas vitaminas, la ganancia por venta de éstas es

$$P = 120x_1 + 180x_2 + 150x_3$$

y como la farmacia quiere competir con los precios de los alimentos F_1 y F_2, de la segunda y tercera columnas de la tabla (6.43), se desprenden las restricciones

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &\leq 10 \quad \text{y} \\ 2x_1 + 3x_2 + x_3 &\leq 12. \end{aligned}$$

Así, lo que se necesita resolver es:

$$\begin{array}{ll} \text{maximizar} & P = 120x_1 + 180x_2 + 150x_3 \\ \text{sujeto a} & \left\{ \begin{array}{ccccccc} x_1 & + & 2x_2 & + & 3x_3 & \leq & 10 \\ 2x_1 & + & 3x_2 & + & x_3 & \leq & 12 \end{array} \right. \\ & x_1, x_2, x_3 \geq 0 \end{array} \tag{6.45}$$

Antes de resolver este problema notemos que la forma matricial de (6.45) es

$$\begin{array}{ll} \text{maximizar} & P = \vec{c}^{\,t}\vec{x} \\ \text{sujeto a} & A\vec{x} \leq \vec{b} \\ & \vec{x} \geq \vec{0}, \vec{b} \geq \vec{0}_{\mathbb{R}^2} \end{array}$$

donde

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}, \vec{b} = \begin{bmatrix} 10 \\ 12 \end{bmatrix} \text{ y } \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix};$$

mientras que la forma matricial de (6.44), del ejemplo 6.33 es

$$\begin{array}{ll} \text{minimizar} & C=\vec{b}^t\vec{y} \\ \text{sujeto a} & A^t\vec{y}\geq\vec{c} \\ & \vec{y}\geq\vec{0}_{\mathbb{R}^3} \end{array}$$

Se dice entonces que cada uno de estos problemas es dual del otro. Ahora resolvamos por el método simplex el problema (6.45): el sistema de igualdades correspondiente al incorporar las variables de holgura es

$$\begin{array}{rcrcrcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & + & y_1 & & & = & 10 \\ 2x_1 & + & 3x_2 & + & x_3 & & & + & y_2 & = & 12 \end{array}$$

y al construir la tabla inicial y aplicar el método simplex tenemos

	x_1	x_2	x_3	y_1	y_2	
y_1	1	2	3	1	0	10
y_2	2	3	1	0	1	12
P	-120	-180	-150	0	0	0

$\sim$

	x_1	x_2	x_3	y_1	y_2	
y_1	$-1/3$	0	$7/3$	1	$-2/3$	2
x_2	$2/3$	1	$1/3$	0	$1/3$	4
P	0	0	-90	0	60	720

$\sim$

	x_1	x_2	x_3	y_1	y_2	
x_3	$-1/7$	0	1	$3/7$	$-2/7$	$6/7$
x_2	$5/7$	1	0	$-1/7$	$3/7$	$26/7$
P	$-90/7$	0	0	$270/7$	$240/7$	$5589/7$

$\sim$

	x_1	x_2	x_3	y_1	y_2	
x_3	0	$1/5$	1	$2/5$	$-1/5$	$8/5$
x_1	1	$7/5$	0	$-1/5$	$3/5$	$26/5$
P	0	18	0	36	42	864

Por tanto, la solución óptima es $P=864$ y se alcanza en $x_1=26/5$, $x_2=0$ y $x_3=8/5$; donde observamos que los valores en la fila objetivo por debajo de las columnas etiquetadas con variables de holgura del problema (6.45), en la tabla final, son precisamente los valores donde se alcanza el mínimo del problema dual (6.44): $y_1=36$, $y_2=42$; además, $P=864=C$ en los dos problemas. Esta no es una simple coincidencia y se cumple para todo par de problemas duales.

Definición 6.2 *Sea el problema estándar de programación lineal*

$$\begin{array}{ll} \textit{maximizar} & P=\vec{c}^{\,t}\vec{x} \\ \textit{sujeto a} & A\vec{x}\leq\vec{b} \\ & \vec{x}\geq\vec{0}_{\mathbb{R}^n},\ \vec{b}\geq\vec{0}_{\mathbb{R}^m} \end{array} \qquad (6.46)$$

A

$$\begin{array}{ll} \textit{minimizar} & C=\vec{b}^t\vec{y} \\ \textit{sujeto a} & A^t\vec{y}\geq\vec{c} \\ & \vec{y}\geq\vec{0}_{\mathbb{R}^m} \end{array} \qquad (6.47)$$

donde $\vec{y}$ es el vector columna que contiene las variables de holgura de (6.46), se le llama problema ***dual*** *de (6.46) y viceversa. Se acostumbra decir de acuerdo a su orden de aparición, problema primal y problema dual, respectivamente, para hacer énfasis en dicho orden.*

Teorema 6.4 (De dualidad) *Si uno de los dos problemas duales (6.46), (6.47) alcanza solución óptima, el otro también y ambos valores óptimos son iguales; es decir,* $P = C$*. Además, si se resuelve el problema (6.46) y el valor máximo* P *se alcanza en la solución factible* $[\vec{x}^t \; \vec{y}^t]$*, entonces el valor óptimo de su problema dual (6.47) se alcanza en* $\vec{y}^t = [\, y_1 \quad y_2 \quad \cdots \quad y_m \,]$ *donde cada* y_i *toma el valor que está por debajo de esta variable, en la última fila, en la tabla final del método simplex para resolver (6.46). Se tiene un resultado análogo si se resuelve (6.47).*

► **Ejemplo 6.34** Resolver el ejemplo (6.44) por medio del método simplex y comprobar para este caso el teorema de dualidad.◄

Solución Transformamos las restricciones (6.44) al sistema de desigualdades equivalente

$$\begin{array}{rcrcr} -y_1 & - & 2y_2 & \leq & -120 \\ -2y_1 & - & 3y_2 & \leq & -180 \\ -3y_1 & - & y_2 & \leq & -150 \end{array}$$

e incorporamos las variables de holgura $x_1, x_2, x_3 \geq 0$ para obtener el sistema

$$\begin{array}{rcrcccccr} -y_1 & - & 2y_2 & + & x_1 & & & = & -120 \\ -2y_1 & - & 3y_2 & & & + & x_2 & = & -180 \\ -3y_1 & - & y_2 & & & & +x_3 & = & -150 \end{array}$$

construimos la matriz (6.37) para este caso y aplicamos la fase I

$$\begin{bmatrix} -1 & -2 & 1 & 0 & 0 & -120 \\ -2 & -3 & 0 & 1 & 0 & -180 \\ -3 & -1 & 0 & 0 & 1 & -150 \\ 10 & 12 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 5 & 0 & 1 & 0 & -2 & 180 \\ 7 & 0 & 0 & 1 & -3 & 270 \\ 3 & 1 & 0 & 0 & -1 & 150 \\ -26 & 0 & 0 & 0 & 12 & -1\,800 \end{bmatrix}$$

para obtener la tabla inicial correspondiente y ejecutar la fase II:

	y_1	y_2	x_1	x_2	x_3	
x_1	5	0	1	0	-2	180
x_2	7	0	0	1	-3	270
y_2	3	1	0	0	-1	150
$-C$	-26	0	0	0	12	$-1\,800$

$\sim$

	y_1	y_2	x_1	x_2	x_3	
y_1	1	0	$1/5$	0	$-2/5$	36
x_2	0	0	$-7/5$	1	$-1/5$	18
y_2	0	1	$-3/5$	0	$1/5$	42
$-C$	0	0	$26/5$	0	$8/5$	-864

Efectivamente, de acuerdo a la tabla final, $x_1 = 26/5$, $x_2 = 0, x_3 = 8/5$ y $C = 864 = P$ para $y_1 = 36$ y $y_2 = 42$, como afirma el teorema 6.4 y como vimos en el ejemplo 6.33 y la discusión que le sigue. ✔

6.5 Teoría de juegos

Requisitos: *Operaciones con matrices y programación lineal (para el apartado 6.6.4): 1.1.1 a 1.1.3 y 6.4.*

6.5.1 Juegos estrictamente determinados y puntos silla

Supongamos que dos personas ejecutan repetidas veces el siguiente juego:

- La persona P elige la cara de una moneda, la persona Q intenta adivinar la cara que escogió P.
- P paga a Q \$4.00, si P elige sol y Q adivina.
- P paga a Q \$3.00, si P elige águila y Q adivina.
- Q paga a P \$5.00, si Q no adivina.

Observamos primero que éste es un juego **de suma cero**, ya que lo que gana uno de los jugadores lo pierde el segundo. Por otro lado, podemos mostrar la ganancia (o pérdida) que tiene en una ejecución dada del juego P y, por tanto, la pérdida (o ganancia) que tiene el jugador Q, en la siguiente matriz de pagos (ganancias con números positivos y pérdidas con números negativos):

$$\begin{array}{cc} & \overbrace{\begin{array}{cc}\text{A} & \text{S}\end{array}}^{Q} \\ P\left\{\begin{array}{c}\text{A}\\ \text{S}\end{array}\right. & \left[\begin{array}{rr} -3 & 5 \\ 5 & -4 \end{array}\right]\end{array}$$

donde las filas representan las elecciones que hace P y las columnas las que hace Q.

En general, si $A = [a_{ij}]$ es una matriz de tamaño $m \times n$, entonces se puede realizar un **juego matricial** de suma cero entre dos jugadores P y Q, el cual consiste en que P elige una fila de A, sin que Q sepa cuál fila eligió, y entonces Q hace lo propio, de manera independiente, con una columna. El elemento a_{ij} de A que se encuentra en la fila i y columna j que escogieron P y Q, respectivamente, es el pago que hace Q a P si $a_{ij} \geq 0$ y el pago que hace P a Q si $a_{ij} \leq 0$. En este caso P tiene m posibles elecciones y Q tiene n posibles elecciones cada vez que se repita el juego. Supondremos en toda esta sección, aunque no se haga explícito, que *todo juego será de suma cero*.

▶ **Ejemplo 6.35** Consideremos el siguiente juego de suma cero entre dos jugadores P y Q con matriz de pagos para P (P juega por filas) dada por

$$A = \left[\begin{array}{rr} 3 & 5 \\ -2 & -3 \end{array}\right].$$

Es evidente que el jugador P siempre elegirá la primera fila, pues con esta elección por lo menos gana \$3.00. Mientras que el jugador Q siempre elegirá la primera columna porque así minimiza su pérdida, lo más que pierde son \$2.00. Este juego es desfavorable para Q y favorable para P. Ambos jugadores han usado la mejor estrategia posible y el valor del juego es entonces \$3.00.◀

Note que en el ejemplo precedente el elemento en la primera fila y primera columna resulta ser el mayor de la columna donde se encuentra y el menor de la fila donde él está. En general, un juego matricial es **estrictamente determinado** si y sólo si la matriz de pagos contiene un elemento a_{ij} que es el menor de la fila i y el mayor de la columna j; en tal caso **el valor del juego** es dicho elemento a_{ij} y se le llama **punto silla**. Así, el juego de las monedas al inicio de este segmento que tiene por matriz de pagos a

$$M = \left[\begin{array}{rr} -3 & 5 \\ 5 & -4 \end{array}\right]$$

no es estrictamente determinado. Los juegos realmente interesantes son como éstos, los no estrictamente determinados y a su estudio nos abocamos en lo que resta de esta sección.

6.5.2 Estrategias y pagos esperados

Regresemos al juego de las monedas con la matriz de pagos precedente. ¿Qué estrategia debe seguir entonces cada jugador? Evidentemente, la que a la larga maximice la ganancia para uno y minimice la pérdida para el otro. Si P sigue una estrategia, por ejemplo, elegir siempre la primera fila, entonces Q se dará cuenta de esto y contrarrestará, en perjuicio de P, eligiendo siempre la primera columna. Si P comienza a elegir con mucha frecuencia la segunda fila, entonces Q elegirá constantemente la segunda columna. Parece ser que no conviene a ninguno de los dos jugadores seguir algún tipo de patrón, sino más bien hacer en cada jugada una elección racional al azar.

Supongamos que P y Q juegan un juego de suma cero no estrictamente determinado con matriz de pagos

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

donde P elige por filas y Q por columnas y que P y Q distribuyen en forma aleatoria las elecciones de filas y columnas como sigue:

1. P elige la primera fila con una proporción fija p_1 y la segunda fila con una proporción $p_2 = 1 - p_1$.
2. Q elige la primera columna con una proporción fija q_1 y la segunda columna con una proporción $q_2 = 1 - q_1$.

Lo anterior significa que la probabilidad de que P elija la primera fila es p_1, que elija la segunda es $p_2 = 1 - p_1$; mientras que la probabilidad de que Q elija la primera columna es q_1 y de que elija la segunda columna es $q_2 = 1 - q_1$. De esta manera, aunque las probabilidades son fijas, es imposible que uno de los jugadores pueda adivinar en una jugada dada qué seleccionará el contrincante. Por ejemplo, si $p_1 = 3/4$ y $q_1 = 2/3$, entonces el jugador P decidirá por la primera fila tres cuartas partes de las veces que juegue y la fila dos, una cuarta parte de estas veces; mientras que el jugador Q optará dos terceras partes por la primera columna y una tercera parte por la segunda columna. Supongamos que éste es el caso; es decir, P y Q asignan las probabilidades de los dos incisos anteriores para las elecciones de filas y columnas, respectivamente; se dice entonces que los vectores de probabilidad $\vec{p} = (p_1, p_2)$ y $\vec{q} = (q_1, q_2)$ son **estrategias** para sendos jugadores. Así, después de varias ejecuciones del juego matricial, el jugador P ha ganado en promedio

$$\begin{aligned} E(\vec{p}, \vec{q}) &= p_1 a_{11} q_1 + p_1 a_{12} q_2 + p_2 a_{21} q_1 + p_2 a_{22} q_2 \\ &= \begin{bmatrix} p_1 & p_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \end{aligned}$$

ya que $p_i q_j$ es claramente la proporción de las veces que se ejecute el juego en las que P seleccionará la fila i y Q la columna j; con lo que se obtendrá el valor a_{ij} en una proporción $p_i q_j$ de las veces que se ejecuta el juego. Al valor númerico $E(\vec{p}, \vec{q})$ se le llama el valor esperado del juego para las estrategias $\vec{p}$ y $\vec{q}$.

Definición 6.3 *Sean P y Q dos contrincantes que juegan un juego matricial no estrictamente determinado de suma cero con matriz de pagos $A = [a_{ij}]$ de tamaño $m \times n$. Se supone que el jugador P elige la fila i con probabilidad p_i y el jugador Q elige la columna j con probabilidad q_j; con $0 \leq p_i \leq 1$, $0 \leq q_j \leq 1$ para todo $i = 1, \ldots m$, para todo $j = 1, \ldots, n$; $\sum_{i=1}^{m} p_i = 1$ y $\sum_{j=1}^{n} q_j = 1$. Si $\vec{p} = (p_1, p_2, \ldots, p_m) = [\ p_1 \quad p_2 \quad \cdots \quad p_m\ ]^t$ y $\vec{q} = (q_1, q_2, \ldots, q_m) = [\ q_1 \quad q_2 \quad \cdots \quad q_n\ ]^t$, se define*

$$E(\vec{p}, \vec{q}) = \vec{p}^{\,t} A \vec{q}$$

*y se llama el **valor esperado del juego** para la **estrategia** $\vec{p}$ del jugador P y para la **estrategia** $\vec{q}$ del jugador Q.*

Intuitivamente $E(\vec{p}, \vec{q})$ es el valor que en promedio gana (o pierde) el jugador P (pierde o gana el jugador Q) después de realizar un considerable número de juegos.

► **Ejemplo 6.36** Si $A = \begin{bmatrix} -3 & 5 \\ 5 & -4 \end{bmatrix}$, $\vec{p} = (2/3, 1/3)$ y $\vec{q} = (1/2, 1/2)$,

$$\begin{aligned} E(\vec{p}, \vec{q}) &= \begin{bmatrix} 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} -3 & 5 \\ 5 & -4 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \\ &= \frac{5}{6}\,. \end{aligned}$$

Entonces, a la larga, el jugador P tendrá en promedio una ganancia de $5/6$ (en unidades de lo que estén apostando) y el jugador Q tendrá en promedio esta pérdida.◄

En el ejemplo anterior el juego es favorable para P y desfavorable para Q. En general, ¿cuál es la mejor estrategia que deben seguir dos jugadores en un juego matricial no determinado estrictamente? Es evidente que P debe elegir $\vec{p}$ para maximizar su ganancia promedio y Q debe escoger $\vec{q}$, de tal suerte que minimice su pérdida promedio. Parece una labor sumamente complicada que tanto P como Q encuentren sendos vectores de probabilidad $\vec{p}$ y $\vec{q}$, de tal suerte que ambos logren ese objetivo. En el siglo pasado John von Neumann, uno de los grandes matemáticos que han existido, demostró que sí existen vectores de probabilidad $\vec{p}$ y $\vec{q}$ para los cuales se alcanzan esos valores óptimos.

Teorema 6.5 **(Minimax)** *Si $A = [a_{ij}]$ una matriz de pagos de un juego de suma cero con dos contrincantes P y Q, entonces existen $v \in \mathbb{R}$, $\vec{p}_0$ y $\vec{q}_0$, estrategias para sendos jugadores, tales que:*

1. *$v \leq E(\vec{p}_0, \vec{q})$ para toda estrategia $\vec{q}$ de Q.*
2. *$E(\vec{p}, \vec{q}_0) \leq v$ para toda estrategia $\vec{p}$ de P.*

El teorema precedente tiene implicaciones sumamente importantes y bien vale la pena ahondar un poco más en las que nos interesan en esta sección:

- El primer inciso de este teorema significa que existe un valor mínimo garantizado, v, como ganancia promedio para P con la estrategia $\vec{p}_0$ sin importar la estrategia $\vec{q}$ que utilice el jugador Q.
- En contraparte, existe una estrategia $\vec{q}_0$ para el jugador Q, tal que su máxima pérdida en promedio es v, independientemente de la estrategia que use P.

- Si ambos jugadores utilizan las estrategias $\vec{p}_0$ y $\vec{q}_0$ de este teorema, entonces $E(\vec{p}_0,\vec{q}_0) \le v \le E(\vec{p}_0,\vec{q})$ y, por tanto,

$$E(\vec{p}_0,\vec{q}_0) = v.$$

❍ **Nota 6.3** Debido a que el jugador Q desea minimizar su máxima pérdida, a la proposición 6.5 se le conoce como *teorema Minimax*.

Definición 6.4 *Los vectores de probabilidad $\vec{p}_0$ y $\vec{q}_0$ del teorema 6.5 se llaman **estrategias óptimas** de los jugadores P y Q, respectivamente, para el juego matricial $A = [a_{ij}]$; mientras que a $E(\vec{p}_0,\vec{q}_0)$ se le conoce como el **valor esperado del juego**.*

El teorema 6.5 sólo postula la existencia de estrategias óptimas y el valor esperado; sin embargo, no dice cómo encontrar esos valores. Para el caso de una matriz de pagos 2×2 es muy sencillo dar un algoritmo para ello; pero el caso general es más complicado y lo veremos más adelante. Antes de hacer plausible el resultado para matrices 2×2 necesitamos de unos cuantos preliminares que hacemos patentes a continuación; la demostración es sencilla y se deja de ejercicio al lector.

Proposición 6.10 *Sea $A = [a_{ij}]$ una matriz de pagos de un juego de suma cero que no está estrictamente determinado para un par de contrincantes P y Q, $v \in \mathbb{R}$ y $\vec{p}_0$, $\vec{q}_0$ estrategias de sendos oponentes. Entonces:*

1. *$\vec{p}_0^{\,t}A \ge v$ si y sólo si $E(\vec{p}_0,\vec{q}) = \vec{p}_0^{\,t}A\vec{q} \ge v$ para toda estrategia $\vec{q}$ de Q.*
2. *$\vec{q}_0^{\,t}A \le v$ si y sólo si $E(\vec{p},\vec{q}_0) = \vec{p}^{\,t}A\vec{q}_0 \le v$ para toda estrategia $\vec{p}$ de Q.*

6.5.3 Estrategias óptimas y valor esperado para juegos matriciales con matriz de pagos 2×2

Ahora supongamos que tenemos un par de oponentes, P y Q, que ejecutan un juego de suma cero que no es estrictamente determinado con matriz de pagos

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

No es difícil probar, por ser el juego no estrictamente determinado, que $a_{11} + a_{22} - a_{12} - a_{21} \neq 0$. Sea $\vec{p} = (p, 1-p)$ una estrategia cualquiera de P. En la figura 6-18*(a)* se bosquejan las gráficas de las rectas $E_1(p) = a_{11}p + a_{21}(1-p)$ y $E_2(p) = a_{12}p + (1-p)a_{22}$ ambas como funciones de p. Sea v el valor esperado del juego, el cual existe por el teorema 6.5, entonces, por el primer inciso de la proposición anterior, se debe tener

$$\begin{aligned} a_{11}p + a_{21}(1-p) &\ge v \qquad \text{y} \\ a_{12}p + (1-p)a_{22} &\ge v. \end{aligned}$$

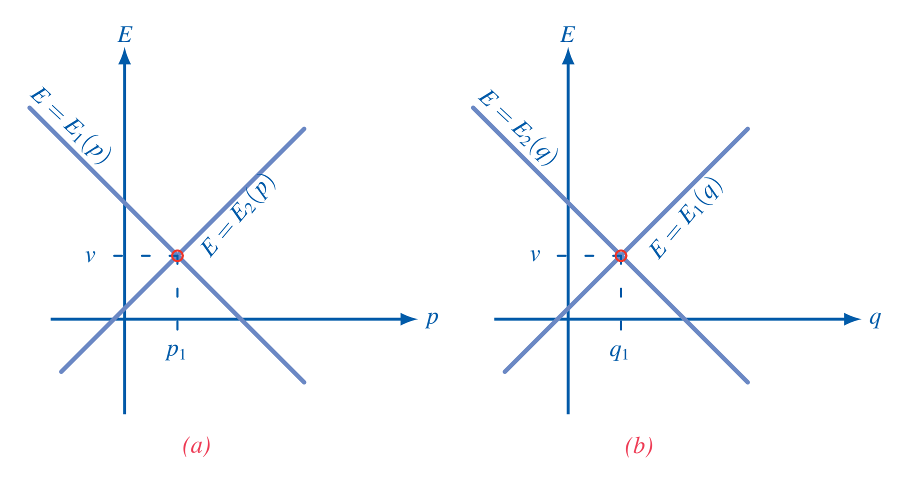


Figura 6-18 •

Al observar la figura 6-18*(a)*, se concluye que el único valor de p para el que se cumplen ambas desigualdades es $p = p_1$, donde se intersecan las rectas $E = E_1(p)$ y $E = E_2(p)$ y que v es precisamente la ordenada que tienen en común ambas rectas para el valor p_1. Unos cálculos simples producen

$$p_1 = \frac{a_{22} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}} \quad \text{y}$$

$$v = \frac{a_{11}a_{22} - a_{21}a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}} \ .$$

Por otra parte, la figura 6-18*(b)* contiene el bosquejo de las rectas $F_1(q) = a_{11}q + a_{12}(1-q)$ y $F_2(q) = a_{21}q + (1-q)a_{22}$, ambas en función de todos los posibles valores q para las estrategias $(q, 1-q)$ de Q. Del segundo inciso de la proposición 6.10 se debe cumplir

$$\begin{aligned} a_{11}q + a_{12}(1-q) &\leq v \\ a_{21}q + (1-q)a_{22} &\leq v \end{aligned}$$

Nuevamente, del bosquejo en la figura 6-18*(b)* se deduce que el único valor $q = q_1$ para el que se cumplen ambas restricciones es la abscisa donde se intersecan las rectas $E = F_1(q)$ y $E = F_2(q)$ y que v debe ser la ordenada que tienen ambas rectas para el valor q_1. Desarrollando el correspondiente cálculo se tiene

$$q_1 = \frac{a_{22} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}} \quad \text{y}$$

$$v = \frac{a_{11}a_{22} - a_{21}a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}} \ .$$

En resumen: las estrategias óptimas $\vec{p}_0$, $\vec{q}_0$ para sendos jugadores P y Q, en un juego de suma cero no estrictamente determinado y con una matriz de pagos $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, el valor esperado del juego están dados por

$$\vec{p}_0 = \begin{bmatrix} \dfrac{a_{22} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}} \\ \\ \dfrac{a_{11} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}} \end{bmatrix}, \ \vec{q}_0 = \begin{bmatrix} \dfrac{a_{22} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}} \\ \\ \dfrac{a_{11} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}} \end{bmatrix}, \qquad (6.48)$$

$$E(p_0, q_0) = \frac{a_{11}a_{22} - a_{21}a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}}, \qquad (6.49)$$

respectivamente.

▶ **Ejemplo 6.37** Calcular las estrategias óptimas del juego de las monedas dado al inicio de esta sección; es decir, para la matriz de pagos

$$\begin{bmatrix} -3 & 5 \\ 5 & -4 \end{bmatrix}. ◀$$

Solución Sustituyendo $a_{11} = -3$, $a_{12} = 5$, $a_{21} = 5$ y $a_{22} = -4$ en (6.48) y (6.49), se obtiene

$$\vec{p}_0 = (9/17, 8/17),$$
$$\vec{q}_0 = (9/17, 8/17) \text{ y}$$
$$E(\vec{p}_0, \vec{q}_0) = 13/17. \quad ✔$$

▶ **Ejemplo 6.38** Resolver el ejemplo precedente, pero con la matriz de pagos

$$\begin{matrix} & & \overbrace{\begin{matrix} \text{A} & \text{S} \end{matrix}}^{Q} \\ P \left\{ \begin{matrix} \text{A} \\ \text{S} \end{matrix} \right. & & \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \end{matrix} ◀$$

Solución Nuevamente al utilizar las fórmulas (6.48) y (6.49),

$$\vec{p}_0 = (1/2, 1/2),$$
$$\vec{q}_0 = (1/2, 1/2) \text{ y}$$
$$E(\vec{p}_0, \vec{q}_0) = 0. \quad ✔$$

Esto produce un juego **justo**, como suele llamarse a todo juego con valor esperado cero. Es claro que el lector pudo, sin tener que hacer un sólo cálculo, dar la misma respuesta por pura intuición, ¿o no?

▶ **Ejemplo 6.39** **(Un juego de política)** En un municipio existen dos candidatos a la presidencia del mismo, el candidato P y el candidato Q. A lo largo de toda la campaña electoral se han propuesto tratar dos asuntos claves, los relacionados con el propio municipio (M) y los relacionados con el estado al que pertenece ese municipio (E). La siguiente matriz contiene los puntos asignados a cada estrategia de los candidatos.

$$\begin{matrix} & & \overbrace{\begin{matrix} \text{M} & \text{E} \end{matrix}}^{Q} \\ P \left\{ \begin{matrix} \text{M} \\ \text{E} \end{matrix} \right. & & \begin{bmatrix} 3 & -2 \\ -1 & 4 \end{bmatrix} \end{matrix}$$

Si el candidato P trata con mucha frecuencia asuntos estatales, el candidato Q puede contrarrestar la campaña de P tratando asuntos municipales. ¿En qué proporciones deben tratar cada uno de los asuntos para encontrar una estrategia óptima si se supone, además, que los puntos que gana uno los pierde el otro? ◀

Solución Es claro que este juego no es estrictamente determinado. Si $\vec{p}_0$ es la estrategia óptima para el candidato P y $\vec{q}_0$ es la mejor estrategia para el candidato Q, al aplicar las fórmulas (6.48) y (6.49) obtenemos

$$\vec{p}_0 = (1/2, 1/2),$$
$$\vec{q}_0 = (3/5, 2/5) \quad \text{y}$$
$$E(\vec{p}_0, \vec{q}_0) = 1.$$

De esta manera el candidato P debe dedicar la mitad de la campaña a los asuntos municipales y la otra mitad a los estatales; el candidato Q tiene por estrategia óptima dedicar $3/5$ partes de la campaña a los asuntos municipales y $2/5$ partes a los estatales. Pero a pesar de ello a la larga el candidato P estará al menos un punto por arriba de su contrincante al final de la campaña. ✔

▶ **Ejemplo 6.40** **(Decisión quirúrgica)** Un paciente posiblemente tenga cierto tumor que puede ser maligno o no. La única opción para tratar esta enfermedad es por medio de una intervención quirúrgica. Si el tumor es maligno y no es extirpado, entonces el paciente tiene una esperanza de vida de 4 años. Si se le practica la operación y el tumor es maligno, tiene una esperanza de vida de 22 años. Debido a efectos secundarios que provoca la intervención quirúrgica, si el tumor no es maligno y el paciente es operado, su esperanza de vida es de 26 años y si el tumor no es maligno y el paciente no es operado, su esperanza de vida es de 32 años. Si la probabilidad de que el tumor que padece el paciente sea maligno es q_1, determinar el umbral de q_1 a partir del cual es recomendable la intervención quirúrgica.◀

Solución Para dar un criterio de decisión al paciente, supongamos que este es un juego de suma cero donde el jugador que juega por renglones es el paciente (P) y el que juega las columnas es la propia naturaleza (N). En la primera fila de la matriz de pagos vamos a considerar la decisión del paciente de ser intervenido quirúrgicamente (I) y la segunda fila la de no serlo (NI); mientras que la primera columna corresponde a que el tumor es maligno (M) y la segunda a que el tumor no lo es (NM). Entonces la matriz de pagos es

$$\begin{array}{cc} & \overbrace{\begin{array}{cc}\text{M} & \text{NM}\end{array}}^{N} \\ P\left\{\begin{array}{c}\text{I}\\ \text{NI}\end{array}\right. & \left[\begin{array}{cc} 22 & 26 \\ 4 & 32 \end{array}\right]\end{array}$$

Es evidente que este problema es estrictamente determinado con valor del juego 22. Sin embargo, convendría al paciente decidir no ser operado, ya que de no ser maligno el tumor la operación reduce la esperanza natural de vida en 6 años. Si decide ser operado, entonces $\vec{p} = (1, 0)$ y, por tanto,

$$\begin{aligned} E_1(\vec{p}, \vec{q}) &= \left[\begin{array}{cc} 1 & 0 \end{array}\right] \left[\begin{array}{cc} 22 & 26 \\ 4 & 32 \end{array}\right] \left[\begin{array}{c} q_1 \\ 1 - q_1 \end{array}\right] \\ &= -4q_1 + 26 \end{aligned}$$

y si decide no operarse $\vec{p} = (0, 1)$ y

$$\begin{aligned} E_2(\vec{p}, \vec{q}) &= \left[\begin{array}{cc} 0 & 1 \end{array}\right] \left[\begin{array}{cc} 22 & 26 \\ 4 & 32 \end{array}\right] \left[\begin{array}{c} q_1 \\ 1 - q_1 \end{array}\right] \\ &= -28q_1 + 32. \end{aligned}$$

Así, sería conveniente para el paciente decidir operarse si $E_1(\vec{p}, \vec{q}) > E_2(\vec{p}, \vec{q})$; esto es, si

$$-4q_1 + 26 > -28q_1 + 32$$

lo cual implica

$$q_1 > \frac{1}{4}.$$

Es decir, el paciente puede decidir operarse si la probabilidad de que el tumor sea maligno es mayor a 0.25. ✔

¿Usted qué haría? Si la probabilidad de que el tumor sea maligno es, por ejemplo, $q_1 = 0.20$, entonces el paciente, de acuerdo a este criterio, no se operaría; sin embargo, aún tiene 20% de posibilidades de que el tumor sea maligno.

6.5.4 Estrategias óptimas y valor esperado con programación lineal para juegos matriciales con matriz de pagos $m \times n$

Se puede probar que si $A = [a_{ij}]$ es una matriz de pagos de tamaño $m \times n$ para un juego de suma cero entre dos contrincantes P y Q, v es el valor esperado del juego, $\vec{p}_0$ y $\vec{q}_0$ son las estrategias óptimas, $r \in \mathbb{R}$ y $B = [a_{ij} + r]$; entonces, el juego con matriz de pagos B tiene las mismas estrategias óptimas que A y su valor esperado es $v + r$. Por tanto, podemos suponer que si A es una matriz de pagos, las componentes de A son números positivos; pues, de no ser así, se puede sumar un número adecuado r a todas ellas para transformarlas en números positivos, calcular las estrategias óptimas de la matriz resultante y su valor esperado v_1; así las estrategias óptimas serán las mismas para A y su valor esperado será $v = v_1 - r$. Supondremos en lo que sigue que $A = [a_{ij}]$ es una matriz de pagos de tamaño $m \times n$ con componentes positivas.

Recordemos que el objetivo del jugador Q es minimizar su máxima pérdida promedio. Sea $\vec{q} = (q_1, q_2, \ldots, q_n)$ el vector variable de probabilidades para las posibles estrategias del jugador Q. Entonces, Q, por el segundo inciso de la proposición 6.10, debe hallar el mínimo de entre todos los números $v \in \mathbb{R}$, tales que todas las componentes de $A\vec{q}$ son menores o iguales a v; esto es:

$$\begin{array}{ll} \text{Hallar el valor mínimo} & v \\ \text{tal que} & \sum_{j=1}^{n} a_{ij} q_j \leq v \\ & \forall\, i = 1, \ldots, m \end{array} \tag{6.50}$$

Ya que $\vec{q}$ es un vector de probabilidad ($0 \leq q_j \leq 1$ y $q_1 + \cdots + q_n = 1$) se infiere que la solución v de (6.50) es mayor a cero. Dado que minimizar v significa maximizar $1/v$, el problema (6.50) equivale a

$$\begin{array}{ll} \text{Hallar el valor máximo} & 1/v \\ \text{tal que} & \sum_{j=1}^{n} a_{ij} q_j \leq v \\ & \forall\, i = 1, \ldots, m \end{array} \tag{6.51}$$

Sean, para cada $j = 1, \ldots, n$,

$$x_j = \frac{1}{v} q_j \geq 0$$

entonces,

$$\begin{aligned} x_1 + x_2 + \cdots + x_n &= \frac{1}{v} \sum_{j=1}^{n} q_i \\ &= \frac{1}{v} \end{aligned}$$

y

$$v \geq \sum_{j=1}^{n} a_{ij} q_j = v \sum_{j=1}^{m} a_{ij} x_j.$$

Por tanto, el problema (6.51) se traduce al problema estándar de programación lineal

$$\begin{array}{ll} \text{maximizar} & f = x_1 + x_2 + \cdots + x_n \\ \text{sujeto a} & \left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & \leq & 1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & \leq & 1 \\ & \vdots & & \vdots & & \vdots & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & \leq & 1 \end{array} \right. \\ & x_j \geq 0 \quad \forall\, j = 1, 2, \ldots, n \end{array} \tag{6.52}$$

Por otra parte, el jugador P tiene como objetivo maximizar su ganancia mínima promedio. Sea $\vec{p} = (p_1, p_2, \ldots, p_m)$ el vector variable de probabilidades para las posibles estrategias de P. Entonces, del primer inciso de la proposición 6.10, debe encontrar el máximo de los valores v, tales que todas las componentes de la matriz $\vec{p}A$ sean mayores o iguales a v; esto es, resolver el problema.

$$\begin{array}{ll} \text{Hallar el valor máximo} & v \\ \text{tal que} & \sum_{i=1}^{m} a_{ij} p_i \geq v \\ & \forall\, j = 1, \ldots, n \end{array}$$

que equivale (ya que hemos supuesto los $a_{ij} > 0$ y que $\vec{p}$ es un vector de probabilidad, se tiene necesariamente $v > 0$) al problema:

$$\begin{array}{ll} \text{Hallar el valor mínimo} & 1/v \\ \text{tal que} & \sum_{i=1}^{m} a_{ij} p_i \geq v \\ & \forall\, j = 1, \ldots, n \end{array}$$

Sean, para cada $i = 1, \ldots, m$,

$$y_i = \frac{p_i}{v},$$

entonces,

$$\sum_{i=1}^{m} y_i = \frac{1}{v}$$

y

$$v \leq \sum_{i=1}^{m} a_{ij} p_i = v \sum_{i=1}^{m} a_{ij} y_i$$

Así, P tiene que resolver el problema de programación lineal

$$\begin{array}{ll} \text{minimizar} & g = y_1 + y_2 + \cdots + y_m \\ \text{sujeto a} & \left\{ \begin{array}{ccccccccc} a_{11}y_1 & + & a_{21}y_2 & + & \cdots & + & a_{m1}y_m & \geq & 1 \\ a_{12}y_1 & + & a_{22}y_2 & + & \cdots & + & a_{m2}y_m & \geq & 1 \\ & \vdots & & \vdots & & \vdots & & \vdots & \\ a_{1n}y_1 & + & a_{2n}y_2 & + & \cdots & + & a_{mn}y_m & \geq & 1 \end{array} \right. \\ & y_i \geq 0 \quad \forall\, i = 1, 2, \ldots, m \end{array} \tag{6.53}$$

Resulta claro que los problemas de programación lineal (6.52) y (6.53) son duales uno del otro y, por tanto, basta resolver uno para conocer los valores óptimos del otro como lo hicimos en la sección precedente. En resumen:

Método para encontrar estrategias óptimas por medio de programación lineal

Sea A una matriz de pagos $m \times n$ de un juego de suma cero con dos contrincantes P y Q.

1. Si todas las componentes de la matriz A son positivas, entonces:

   (a) El valor esperado para Q está dada por

   $$v = \frac{1}{f_0}$$

   donde f_0 es el valor óptimo del problema estándar de programación lineal

   $$\begin{array}{ll} \text{maximizar} & f = x_1 + x_2 + \cdots + x_n \\ \text{sujeto a} & A\vec{x} \leq 1 \\ & \vec{x} \geq \vec{0}_{\mathbb{R}^n} \end{array} \tag{6.54}$$

   Si en $\vec{x}_0 = (x_1^0, x_2^0, \ldots, x_n^0)$ se alcanza el valor máximo f_0, entonces

   $$\vec{q}_0 = v\vec{x}_0$$

   es la estrategia óptima para Q.

   (b) La estrategia óptima para el jugador P es el vector $\vec{p}_0 = v(y_1, y_2, \ldots, y_m)$, donde los y_i toman los valores que se encuentran en la fila objetivo debajo de cada variable de holgura en la tabla final para resolver el problema (6.54) por el método simplex. El valor esperado para el jugador P es el mismo para el jugador Q; esto es

   $$E(\vec{p}_0, \vec{q}_0) = v$$

2. Si A tiene algunos registros no positivos, se suma a todas sus componentes una constante r suficientemente grande para obtener una matriz B que tenga todas sus componentes positivas. A la matriz B se le aplica el método del inciso anterior; las estrategias óptimas de A y B son las mismas, pero el valor esperado para la matriz A es el valor esperado para la matriz B, menos la constante r.

▶ **Ejemplo 6.41** Resolver el juego matricial

$$A = \begin{bmatrix} -3 & 5 \\ 5 & -4 \end{bmatrix} ◀$$

Solución Sea v el valor esperado para el juego matricial con esta matriz de pagos. Puesto que la matriz A no tiene todas las entradas positivas, sumamos a cada una de sus componentes $r = 5$ para obtener la matriz

$$B = \begin{bmatrix} 2 & 10 \\ 10 & 1 \end{bmatrix}.$$

Sea v_1 el valor esperado para la matriz de pagos B. Resolvamos el problema estándar de programación lineal

$$\text{maximizar} \quad f_1 = x_1 + x_2$$
$$\text{sujeto a} \quad \begin{cases} 2x + 10y \leq 1 \\ 10x + y \leq 1 \end{cases} :$$
$$x_1, x_2 \geq 0$$

	x_1	x_2	y_1	y_2	
y_1	2	10	1	0	1
y_2	10	1	0	1	1
f_1	−1	−1	0	0	0

$\sim$

	x_1	x_2	y_1	y_2	
y_1	0	49/5	1	−1/5	4/5
x_1	1	1/10	0	1/10	1/10
f_1	0	−9/10	0	1/10	1/10

$\sim$

	x_1	x_2	y_1	y_2	
x_2	0	1	5/49	−1/49	4/49
x_1	1	0	−1/98	5/49	9/98
f_1	0	0	9/98	4/49	17/98

Luego

$$f_1 = \frac{17}{98}$$
$$\vec{x}_0 = (9/98, 4/49)$$
$$\vec{y}_0 = (9/98, 4/49)$$

Entonces,

$$v_1 = \frac{1}{f_1} = \frac{98}{17}$$

y, por tanto,

$$\vec{q}_0 = v_1\vec{x}_0 = \frac{98}{17}(9/98, 4/49) = (9/17, 8/17),$$

$$\vec{p}_0 = v_1\vec{y}_0 = \frac{98}{17}(9/98, 4/49) = (9/17, 8/17).$$

Finalmente

$$v = v_1 - r = \frac{98}{17} - 5 = \frac{13}{17};$$

que son los mismos valores que calculamos en el ejemplo 6.37. ✔

▶ **Ejemplo 6.42** La empresa Martín tiene dos compañías, I y II. En promedio paga por ellas anualmente $5 000 000.00 y $15 000 000.00, respectivamente, de impuestos. Para cada una de estas compañías la empresa puede declarar los ingresos reales y pagar los impuestos correspondientes, o falsificar la contabilidad y evadir el pago de impuestos. Hacienda tiene únicamente medios para investigar a una de las dos compañías cada año. Si investiga a una compañía y descubre la defraudación, ésta tiene que pagar lo que corresponde a impuestos, más una multa del doble de lo defraudado. Encontrar la estrategia óptima que debe utilizar Hacienda para maximizar los ingresos por recaudación a esta empresa.◀

Solución Construyamos la matriz de pagos, colocando a Hacienda como el contendiente que juega por filas y a la empresa como el contendiente que juega por columnas. Las opciones por fila y columna para Hacienda y la empresa las listamos a continuación:

- H1: Hacienda investiga a compañía I.
- H2: Hacienda investiga a compañía II.
- P1: La empresa declara los impuestos reales de I y II.
- P2: La empresa declara los impuestos reales de I y evade los de II.
- P3: La empresa declara los impuestos reales de II y evade los de I.
- P4: La empresa evade los impuestos de I de II.

Entonces,

$$\begin{array}{c} \\ \text{H1} \\ \text{H2} \end{array} \begin{array}{c} \begin{array}{cccc} \text{P1} & \text{P2} & \text{P3} & \text{P4} \end{array} \\ \begin{bmatrix} 15 & 5 & 25 & 15 \\ 15 & 35 & 10 & 30 \end{bmatrix} \end{array}$$

es la matriz de pagos correspondiente, donde las entradas están en millones. Sean v el valor esperado, $\vec{p}_0$ y $\vec{q}_0$ las estrategias óptimas para hacienda y la empresa, respectivamente. Resolvamos el problema estándar de programación lineal

$$\begin{array}{ll} \text{maximizar} & f = x_1 + x_2 + x_3 + x_4 \\ \text{sujeto a} & A\vec{x} \leq 1 \\ & \vec{x} \geq \vec{0}_{\mathbb{R}^4} \end{array}$$

donde $\vec{x} = (x_1, \ldots, x_n)$, y A es la matriz de pagos:

	x_1	x_2	x_3	x_4	y_1	y_2	
y_1	15	5	25	15	1	0	1
y_2	15	35	10	30	0	1	1
f	−1	−1	−1	−1	0	0	0

$\sim$

	x_1	x_2	x_3	x_4	y_1	y_2	
x_1	1	1/3	5/3	1	1/15	0	1/15
y_2	0	30	−15	15	−1	1	0
f	0	−2/3	2/3	0	1/15	0	1/15

$\sim$

	x_1	x_2	x_3	x_4	y_1	y_2	
x_1	1	0	11/6	5/6	7/90	−1/90	1/15
x_2	0	1	−1/2	1/2	−1/30	1/30	0
f	0	0	1/3	1/3	2/45	1/45	1/15

Por tanto,

$$f = 1/15$$
$$\vec{x}_0 = (1/15, 0, 0, 0)$$

y la solución para el problema dual correspondiente es

$$\vec{y}_0 = (2/45, 1/45).$$

Así,

$$\begin{aligned} v &= 15, \\ \vec{q}_0 &= 15\vec{x}_0 = (1,0,0,0) \quad \text{y} \\ \vec{p}_0 &= 15\vec{y}_0 = (2/3, 1/3). \end{aligned}$$

Hacienda puede meter tres bolas del mismo tamaño en una urna, dos rojas y una blanca, revolver bien en el interior las bolas y después extraer una. Si la bola es roja, debe investigar a la compañía I y si es blanca a la compañía II; mientras que la mejor estrategia para la empresa es declarar los impuestos reales de ambas compañías. ✔

6.5.5 Filas y columnas recesivas o dominantes

En ocasiones un problema de teoría de juegos se simplifica cuando es posible reducir el tamaño de la matriz de pagos porque tiene alguna fila o columna que es factible eliminar. Por ejemplo, si la matriz de pagos de un juego de dos personas de suma cero es

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 1 & -2 & 1 \\ -1 & 4 & 3 \end{bmatrix},$$

entonces, el jugador que juega por filas va a elegir siempre la primera fila en lugar de la segunda, porque con esta elección tendrá una mayor ganacia y una menor pérdida. De manera similar el jugador cuyas opciones son las columnas, siempre elegirá la primera columna en lugar de la tercera, porque ésta tendrá una menor pérdida y una mayor ganancia.

Definición 6.5 *Sea A una matriz de pagos de un juego de suma cero de dos jugadores.*

1. *Supongamos que existen un par de filas de A, F_i y F_k, tales que*[12] $F_i \geq F_k$. *Entonces se dice que la fila F_i domina a la fila F_k. La F_i es una **fila dominante** y la fila dominada, F_k, es una **fila recesiva**.*
2. *Si existen un par de columnas de A, K_j y K_l, tales que $K_j \leq K_l$, se dice que la columna K_j domina a la columna K_l y que ésta es, entonces, una **columna recesiva** y K_j es una **columna dominante**.*

Por tanto, si F_k y K_l son una fila y una columna recesivas, podemos excluir ambas de la matriz A y trabajar con la matriz B de tamaño $(m-1)\times(n-1)$ que resulta de excluir la fila y columna recesivas de A. Si, además, $\vec{\pi}_0 = (p_1,\ldots,p_{k-1},p_{k+1},\ldots,p_m)$ y $\vec{\chi}_0 = (q_1,\ldots,q_{l-1},q_{l+1},\ldots,q_n)$ son las estrategias óptimas de B, entonces las estrategias óptimas para la matriz de pagos A son

$$\vec{p}_0 = (p_1,\ldots,p_{k-1},0,p_{k+1},\ldots,p_m)$$
$$\vec{q}_0 = (q_1,\ldots,q_{l-1},0,q_{l+1},\ldots,q_n)$$

y el valor esperado es el mismo para ambas matrices. Es conveniente tener esto siempre presente y, de ser posible, eliminar filas y columnas recesivas antes de aplicar el método simplex para trabajar con matrices de tamaño menor.

[12] Recuerde que la notación $U \leq V$ para un par de matrices $U = [u_{ij}]$, $V = [v_{ij}]$ del mismo tamaño, significa $u_{ij} \leq v_{ij}\ \forall\, i, j$.

► **Ejemplo 6.43** Encontrar las estrategias óptimas y el valor esperado de la matriz de pagos

$$A = \begin{bmatrix} -6 & -4 & 2 \\ 2 & -1 & 2 \\ -3 & 4 & 4 \end{bmatrix}. ◄$$

Solución Se tiene $F_2 \geq F_1$, por tanto, la fila F_1 es recesiva (es dominada por F_2). También $K_1 \leq K_3$, por lo que K_3 es una columna recesiva dominada por K_1. Entonces trabajamos con

$$B = \begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix}.$$

Podemos aplicar las fórmulas (6.48) y (6.49) para obtener

$$\vec{\pi}_0 = (7/10, 3/10),$$
$$\vec{\chi}_0 = (1/2, 1/2) \text{ y}$$
$$v = 1/2.$$

Por tanto, las estrategias óptimas para la matriz de pagos original son:

$$\vec{p}_0 = (0, 7/10, 3/10),$$
$$\vec{q}_0 = (1/2, 1/2, 0) \text{ y}$$
$$E(\vec{p}_0, \vec{q}_0) = 1/2. \quad ✔$$

6.6 Cadenas de Markov

Requisitos: *Nociones de probabilidad y del concepto de límite; operaciones con matrices: 1.1.1 a 1.1.3.*

Una ciudad cuenta únicamente con dos empresas de alquiler de autos, A_1 y A_2. Los porcentajes iniciales de la clientela son de 35 % y 65 %, respectivamente. La empresa A_1 ha lanzado una extensa campaña de publicidad para competir con la empresa A_2 y así incrementar el porcentaje de su clientela. El registro de ventas mensuales revela lo siguiente:

(a) 80 % de los clientes de la empresa A_1 regresan a utilizar sus servicios.
(b) 15 % de los clientes de la empresa A_2 cambian a la empresa A_1.

Veamos cómo evolucionan los porcentajes de clientela de estas empresas en forma mensual. Para ello, sea $\vec{p}_t = (p_1(t), p_2(t))$ el vector que contiene las proporciones de clientes de cada empresa en el mes t; es decir, $p_1(t)$ es la proporción de clientes de la empresa A_1 y $p_2(t)$ es la proporción de clientes de la empresa A_2. Esto significa que $p_i(t)$ es la probabilidad, en el periodo t, que un cliente que alquila un auto lo haga en la empresa A_i; de esta manera $\vec{p}(t)$ es un vector de probabilidades y, por tanto, $0 \leq p_i(t) \leq 1$ y $p_1(t) + p_2(t) = 1$. Sabemos que 80 % de los clientes de A_1 mes a mes continúa siendo cliente de A_1 y que 15 % de los clientes de A_2 pasa a ser cliente de A_1, por tanto,

$$\begin{aligned} p_1(1) &= (0.80)(0.35)+(0.15)(0.65) \\ &= 0.3775 \end{aligned}$$

Dado que mes a mes 20 % de los clientes de A_1 pasan a ser clientes de A_2 y 15 % de los clientes de A_2 pasa a ser cliente de A_1, y esta empresa retiene, periodo a periodo, 85 % de su clientela,

$$\begin{aligned} p_2(1) &= (0.20)(0.35)+(0.85)(0.65) \\ &= 0.6225 \end{aligned}$$

Es decir, al cabo del primer mes, los porcentajes de clientes son 37.75 % para la empresa A_1 y 62.25 % para la empresa A_2. Ahora bien, es posible determinar esta información de manera matricial. La siguiente tabla resume la información dada para este problema en los incisos *(a)* y *(b)*:

	A_1	A_2
A_1	0.80	0.15
A_2	0.20	0.85

(6.55)

El significado es: la intersección de la columna etiquetada con el símbolo A_j y la fila etiquetada con la letra A_i contiene la proporción de clientes de la empresa A_j que, mes a mes, pasan a ser clientes de la empresa A_i; esto es, la probabilidad de que un cliente que recibía un servicio en el periodo $t-1$ en la empresa A_j, contrate un servicio en la empresa A_i en el periodo t, para $t = 1, 2, \ldots$. Así, por ejemplo, la intersección de la segunda columna con la primera fila contiene la proporción de clientes que mes a mes pasa de ser cliente de A_2 a ser cliente de A_1; la probabilidad de que un cliente que recibió servicio de la empresa A_2 en un mes dado, sea cliente de A_1 en el siguiente mes. Note que las proporciones (probabilidades) contenidas en esta tabla permanecen constantes mes a mes. Si T es la matriz de probabilidades de esta tabla; esto es,

$$T = \begin{bmatrix} 0.80 & 0.15 \\ 0.20 & 0.85 \end{bmatrix}$$

observamos que

$$\begin{aligned} T\vec{p}_0 &= \begin{bmatrix} 0.80 & 0.15 \\ 0.20 & 0.85 \end{bmatrix}\begin{bmatrix} 0.35 \\ 0.65 \end{bmatrix} \\ &= \begin{bmatrix} 0.3775 \\ 0.6225 \end{bmatrix} \\ &= \vec{p}_1. \end{aligned}$$

Dado que al cabo del primer mes los porcentajes de clientes de las empresas A_1 y A_2 son 37.75 % y 62.25 %, 80 % de clientes de A_1 permanece con A_1 y 15 % de clientes de A_2 pasan a ser clientes de A_1, tenemos

$$\begin{aligned} p_1(2) &= (0.3775)(0.80)+(0.6225)(0.15) \\ &= 0.39538 \end{aligned}$$

Puesto que 20 % de clientes de A_1 pasa a ser cliente de A_2, y A_2 retiene 85 % de su clientela,

$$\begin{aligned} p_2(2) &= (0.3775)(0.20)+(0.6225)(0.85) \\ &= 0.60463 \end{aligned}$$

Esto es

$$
\begin{aligned}
\vec{p}_2 &= \begin{bmatrix} 0.39538 \\ 0.60463 \end{bmatrix} \\
&= \begin{bmatrix} (0.3775)(0.80)+(0.6225)(0.15) \\ (0.3775)(0.20)+(0.6225)(0.85) \end{bmatrix} \\
&= \begin{bmatrix} 0.80 & 0.15 \\ 0.20 & 0.85 \end{bmatrix} \begin{bmatrix} 0.3775 \\ 0.6225 \end{bmatrix} \\
&= T\vec{p}_1 \\
&= T(T\vec{p}_0) \\
&= T^2\vec{p}_0
\end{aligned}
$$

Si continuamos así, podemos ver que en el mes m

$$\vec{p}_m = T^m\vec{p}_0 \,.$$

¿Qué es una cadena de Markov?

En general, se tiene una población Ω y un número finito de características o estados, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, de tal manera que todo individuo posee una y sólo una de las características ε_i o, como se suele decir, todo individuo de Ω se encuentra en uno y sólo un **estado** ε_i, como se ilustra en la figura 6-19*(a)*. La proporción de individuos de la población que se encuentra inicialmente en el estado ε_i es $p_i(0)$, $i = 1, \ldots, n$, es la probabilidad de que un individuo de Ω se encuentre en el estado ε_i. El vector de probabilidad $\vec{p}_0 = (p_1(0), p_2(0), \ldots, p_n(0))$ es la **distribución inicial de la población**. Después de cierto intervalo τ, que llamaremos **periodo**, la población se ha redistribuido, algunos individuos han pasado de un estado ε_j a un estado ε_i; mientras que otros permanecen en el mismo estado (vea la figura 6-19*(b)*). Ahora Ω tiene una distribución de población $\vec{p}_1 = (p_1(1), p_2(1), \ldots, p_n(1))$, que contiene las nuevas proporciones (probabilidades), $p_i(1)$, de los individuos que se encuentran en el estado ε_i, $i = 1, \ldots n$, para el primer periodo. Transcurrido un periodo más, la población nuevamente se ha redistribuido con una distribución de probabilidad $\vec{p}_2 = (p_1(2), p_2(2), \ldots, p_n(2))$ para el segundo periodo y así sucesivamente. Después de m periodos Ω tiene una nueva **distribución de población** $\vec{p}_m = (p_1(m), p_2(m), \ldots, p_n(m))$ que contiene las proporciones de la población que se encuentran en cada estado para el m-ésimo periodo; las probabilidades de que un individuo se encuentre en cada estado para ese periodo [(figura 6-19*(c)*]. Además, se supone que las probabilidades de que se pase del estado ε_j al estado ε_i; esto es, las probabilidades de que un individuo pase del estado ε_j al estado ε_i de un periodo al periodo inmediato posterior, permanecen constantes, periodo a periodo. Es decir, las proporciones de la población que cambian de un estado a otro de un periodo al siguiente permanecen constantes en todo el proceso. Bajo las condiciones precedentes, a la sucesión de distribuciones poblacionales $\{\vec{p}_m\}$ le llamaremos **cadena o proceso de Markov**.

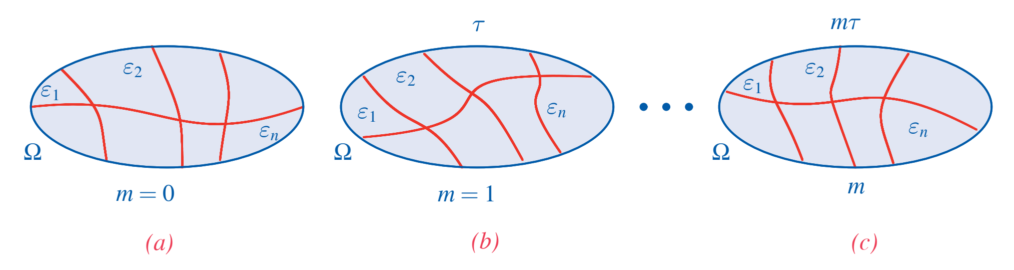


Figura 6-19 •

Matriz de transición

Definición 6.6 *Sea $\{\vec{p}_m\}$ un proceso de Markov, denotemos por a_{ij} la probabilidad de pasar del estado j al estado i. A la matriz*

$$T = [a_{ij}]$$

*se llama **matriz de transición** del proceso de Markov.*

En la siguiente proposición resumimos las propiedades básicas que tienen las matrices de transición, la demostración de cada una de ellas es simple y se deja de ejercicio al lector.

Proposición 6.11 *Sea $T = [a_{ij}]$ la matriz de transición de un proceso de Markov $\{\vec{p}_m\}$. Entonces:*

1. $0 \leq a_{ij} \leq 1\ \forall\, i, j$; *además,* $\sum_i^n a_{ij} = 1$ *para todo j; es decir, las columnas de T son vectores de probabilidad.*
2. *Si $\vec{p} \in \mathbb{R}^n$ es un vector de probabilidad, entonces $T\vec{p}$ también es un vector de probabilidad.*
3. *Si $\vec{p}_0 \in \mathbb{R}^n$ es un vector de probabilidad, entonces*

$$\begin{aligned}\vec{p}_m &= T\vec{p}_{m-1} \quad y\\ \vec{p}_m &= T^m\vec{p}_0\end{aligned}$$

   para todo $m = 1, 2, \ldots$

4. *Si T_1 es una matriz cuyas columnas son vectores de probabilidad, entonces T_1T es también una matriz cuyas columnas son vectores de probabilidad. En particular, T^m es una matriz con sus columnas vectores de probabilidad.*

El objetivo fundamental en los procesos de Markov, es determinar las distribuciones $\vec{p}_m$ que se van obteniendo en cada periodo y su comportamiento a medida que m aumenta. Es decir, calcular los vectores de probabilidad $\vec{p}_m = T\vec{p}_{m-1} = T^m\vec{p}_0$ y determinar si existe una tendencia de estas distribuciones cuando $m \to \infty$. Por ejemplo, en el caso de las empresas de alquiler de autos presentado antes, $\vec{p}_0 = (0.35, 0.65)$, que contiene la probabilidad de que un cliente, al inicio, contrate el servicio con la empresa A_1 (0.35) y de que un cliente contrate con la empresa A_2 (0.65), obtuvimos $\vec{p}_1 = (0.3775, 0.6225)$, cuyas componentes son las probabilidades de los mismos eventos, respectivamente, pero en el periodo 1 (las proporciones de los clientes que contratan con A_1 y los que contratan con A_2 en el periodo $m = 1$). Con este mismo ejemplo y de manera númerica vamos a analizar la segunda parte del objetivo fundamental. La tabla 6.2 contiene los cálculos, hechos en computadora, de los vectores de probabilidad $T\vec{p}_m$, para varios valores de m, con el fin de conjeturar la tendencia de estos valores. Al observarlos parece ser que la distribución $\vec{p}_m$ tiende a $(0.43, 0.57)$ cuando m tiende a infinito. Es decir, aproximadamente 43 % de la clientela contratará con la empresa A_1 y 57 % con la empresa A_2 en cualquier periodo; los porcentajes $p_i(m)$ tienden a estabilizarse para m suficientemente grande. Vamos a establecer la generalización de este resultado particular haciendo patentes algunos preliminares teóricos, cuyas demostraciones son complicadas y se salen de los objetivos del alcance de esta pequeña sección.

m	p1	p2	m	p1	p2
1	0.37750000	0.62250000	12	0.42812451	0.57187549
2	0.39537500	0.60462500	13	0.42828093	0.57171907
3	0.40699375	0.59300625	16	0.42849165	0.57150835
4	0.41454594	0.58545406	20	0.42855719	0.57144281
5	0.41945486	0.58054514	25	0.42856978	0.57143022
6	0.42264566	0.57735434	30	0.42857124	0.57142876
7	0.42471968	0.57528032	36	0.42857141	0.57142859
8	0.42606779	0.57393221	38	0.42857142	0.57142858
9	0.42694406	0.57305594	39	0.42857142	0.57142858
10	0.42751364	0.57248636	40	0.42857143	0.57142857
11	0.42788387	0.57211613	45	0.42857143	0.57142857

Tabla 6-2 •

Definición 6.7 *Si $T = [a_{ij}]$ es la matriz de transición de un proceso de Markov, se dice que T es* ***regular*** *si existe un entero positivo m_0 tal que T^{m_0} tiene todas sus componentes positivas.*

▶ **Ejemplo 6.44** Si $T_1 = \begin{bmatrix} 1 & 1/2 \\ 0 & 1/2 \end{bmatrix}$, es la matriz de un proceso de Markov, entonces, dado que la componente $(T_1^2)_{21} = 0$, se infiere que $(T^m)_{21} = 0$ para todo $m = 1, 2, \ldots$ En cambio $T_2 = \begin{bmatrix} 0 & 1/2 \\ 1 & 1/2 \end{bmatrix}$ satisface

$$T_2^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

por tanto, T_1 no es regular y T_2 sí.◀

Estado estacionario o de equilibrio

Teorema 6.6 *Sea $T = [a_{ij}]$ una matriz cuadrada de orden n que es la matriz de transición de un proceso de Markov $\{\vec{p}_m\}$. Si T es regular, entonces existe un único vector de probabilidad con componentes estrictamente positivas, $\vec{u} \in \mathbb{R}^n$, tal que:*

1. $T\vec{u} = \vec{u}$.
2. *Todas las columnas de T^m tienden al vector $\vec{u}$ cuando m tiende a infinito. Esto es, si*

$$\vec{u} = (u_1, u_2, \ldots, u_n) = [\ u_1 \quad u_2 \quad \cdots \quad u_n\ ]^t$$

entonces,[13] *para cada $j = 1, \ldots, n$,*

$$\lim_{m\to\infty} (T^m)_{ij} = u_j$$

para todo $i = 1, 2, \ldots n$.

3. $$\lim_{m\to\infty} T^m \vec{p}_0 = \vec{u} \tag{6.56}$$

[13]Recuerde que si $A = [a_{ij}]$, utilizamos la notación $(A)_{ij} = a_{ij}$.

La igualdad (6.56) es muy fácil de intuir, pues si $S = [\underbrace{\vec{u} \quad \vec{u} \quad \cdots \quad \vec{u}}_{n}]$ es la matriz que tiene todas sus columnas iguales al vector $\vec{u}$, entonces $T^m \approx S$ para m grande y, por tanto, si $p_0 = (p_1, p_2, \ldots, p_n)$,

$$\begin{aligned} T^m\vec{p}_0 &\approx S\vec{p}_0 \\ &= \begin{bmatrix} u_1p_1 + u_1p_2 + \cdots + u_1p_n \\ u_2p_1 + u_2p_2 + \cdots + u_2p_n \\ \vdots \\ u_np_1 + u_np_2 + \cdots + u_np_n \end{bmatrix} \\ &= \begin{bmatrix} u_1(p_1 + p_2 + \cdots + p_n) \\ u_2(p_1 + p_2 + \cdots + p_n) \\ \vdots \\ u_n(p_1 + p_2 + \cdots + p_n) \end{bmatrix} \\ &= \begin{bmatrix} u_1 \cdot 1 \\ u_2 \cdot 1 \\ \vdots \\ u_n \cdot 1 \end{bmatrix} \\ &= \vec{u}. \end{aligned}$$

Al vector $\vec{u}$ del teorema 6.6 se le llama **vector de distribución del estado estacionario** del proceso de Markov, porque para m grande las distribuciones $\vec{p}_m$ de la cadena de Markov tiene virtualmente sus valores iguales al vector $\vec{u}$. Además, algo que es sumamente importante es que este vector, cuando la matriz de transición es regular, existe y es **independiente** de la distribución inicial $\vec{p}_0$; pues es el vector común que tiene en todas sus columnas la matriz T^m en el límite cuando m tiende a infinito. Resumimos a continuación esta información:

Método para encontrar el vector de distribución del estado estacionario en un proceso regular de Markov

Sea $T = [a_{ij}] \in \mathfrak{M}_{n\times n}$ la matriz de transición de un proceso de Markov. Para hallar el vector del estado estacionario del proceso se procede como a continuación se indica:

1. Se comprueba que la matriz T sea regular.
2. Se resuelve el sistema homogéneo

$$(T - I_n)\vec{u} = \vec{0}_{\mathbb{R}^n}$$

   y de todas las soluciones $\vec{u} = (u_1, u_2, \ldots, u_n)$ se busca aquella para la cual

$$\begin{aligned} u_1 + u_2 + \cdots + u_n = 1\,, \\ u_i \geq 0\; \forall\, i = 1, \ldots, n. \end{aligned}$$

3. El vector de probabilidad $\vec{u}$ encontrado en el inciso anterior es el vector de distribución del estado estacionario del proceso de Markov.

► **Ejemplo 6.45** Hallar el vector de distribución del estado estacionario de la cadena de Markov del caso de las empresas de alquiler de autos presentado al inicio de esta sección. ◄

Solución Para este caso vimos que la matriz de transición del proceso (cfr. (6.55)) es

$$T = \begin{bmatrix} 0.80 & 0.15 \\ 0.20 & 0.85 \end{bmatrix}.$$

Claramente la matriz T es regular. Tenemos que encontrar un vector de probabilidad $\vec{u}$ con todas sus componentes positivas tal que

$$T\vec{u} = \vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

que equivale a

$$(T - I_2)\vec{u} = \vec{0}_{\mathbb{R}^2}.$$

Hay que resolver el sistema homogéneo llevando a forma escalonada:

$$\begin{bmatrix} 0.80-1 & 0.15 \\ 0.20 & 0.85-1 \end{bmatrix} = \begin{bmatrix} -0.2 & 0.15 \\ 0.2 & -0.15 \end{bmatrix}$$
$$\sim \begin{bmatrix} -0.2 & 0.15 \\ 0 & 0 \end{bmatrix}$$

y al hacer sustitución regresiva, obtenemos

$$u_1 = \frac{3}{4}u_2$$

y como además se debe cumplir

$$u_1 + u_2 = 1$$

se tiene

$$\vec{u} = (3/7, 4/7)$$

que redondeando a dos cifras decimales produce

$$\vec{u} \approx (0.43, 0.57).$$

Que es el resultado numérico que encontramos antes calculando los $T\vec{p}_m$ para valores grandes de m por medio de una computadora (cfr. tabla 6.2, pág. 662). ✔

No debe perderse de vista la interpretación que tienen estos casos en particular; es decir, que a la larga, sin importar en qué mes se esté, los porcentajes de la clientela que contrata en la empresa A_1 y A_2 se establecen en 43 % y 57 %, respectivamante. Observe que en la tabla de valores de estos cálculos (tabla 6.2) a partir del mes 12 casi se tienen estas distribuciones.

► **Ejemplo 6.46** Supongamos que la población de cierto país está dividida en clases, de acuerdo con sus ingresos, en clase pobre (P); clase media (M) y clase rica (R). Inicialmente los porcentajes están repartidos de la siguiente manera:

- 50 % clase pobre.
- 40 % clase media.
- 10 % clase rica.

Se supone que en cada periodo de 25 años:

- 15 % de la clase pobre pasa a clase media.
- 0.5 % de la clase pobre pasa a clase rica.
- 15 % de la clase media pasa a clase pobre.
- 5 % de la clase media pasa a la clase rica.
- 3 % de la clase rica pasa a la clase pobre.
- 10 % de la clase rica pasa a clase media.

Hallar la distribución del estado estacionario del problema si es que existe y comparar con la distribución inicial.◀

Solución Formemos una tabla donde podamos registrar las probabilidades del paso de un estado al otro:

	P	M	R
P	0.845	0.150	0.030
M	0.150	0.800	0.100
R	0.005	0.050	0.870

Por tanto, la matriz de transición es

$$T = \begin{bmatrix} 0.845 & 0.150 & 0.030 \\ 0.150 & 0.800 & 0.100 \\ 0.005 & 0.050 & 0.870 \end{bmatrix}.$$

Resolvamos el sistema homogéneo

$$[T - I_3] = \vec{0}_{\mathbb{R}^3}$$

llevando a forma escalonada:

$$\begin{bmatrix} 0.845-1 & 0.150 & 0.030 \\ 0.150 & 0.800-1 & 0.100 \\ 0.005 & 0.050 & 0.870-1 \end{bmatrix} = \begin{bmatrix} -0.155 & 0.150 & 0.030 \\ 0.150 & -0.200 & 0.100 \\ 0.005 & 0.05 & -0.130 \end{bmatrix}$$

$$\sim \begin{bmatrix} 0.005 & 0.05 & -0.130 \\ 0.150 & -0.200 & 0.100 \\ -0.155 & 0.150 & 0.030 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 10 & -26 \\ 0 & -1.700 & 4 \\ 0 & 1.70 & -4 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 10 & -26 \\ 0 & -1.700 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

resolviendo por sustitución regresiva obtenemos

$$\vec{u} = \begin{bmatrix} \frac{42}{17}u_3 \\ \frac{40}{17}u_3 \\ u_3 \end{bmatrix};$$

como también se debe tener

$$u_1 + u_2 + u_3 = 1$$

se deduce

$$\vec{u} = \begin{bmatrix} 42/99 \\ 40/99 \\ 17/99 \end{bmatrix}$$

Así, los porcentajes de las poblaciones clase pobre, media y rica se estabilizan a la larga en aproximadamente 42.424 %, 40.404 % y 17.172 %. La pobreza disminuye sólo en 7.576 %, la clase media aumenta en únicamente 0.404 %, y la clase rica aumenta 7.172 %. ✔

6.7 Sistemas lineales de ecuaciones diferenciales

Requisitos: *Nociones de ecuaciones diferenciales; valores propios y vectores propios; diagonalización: 5.3.1 y 5.3.2.*

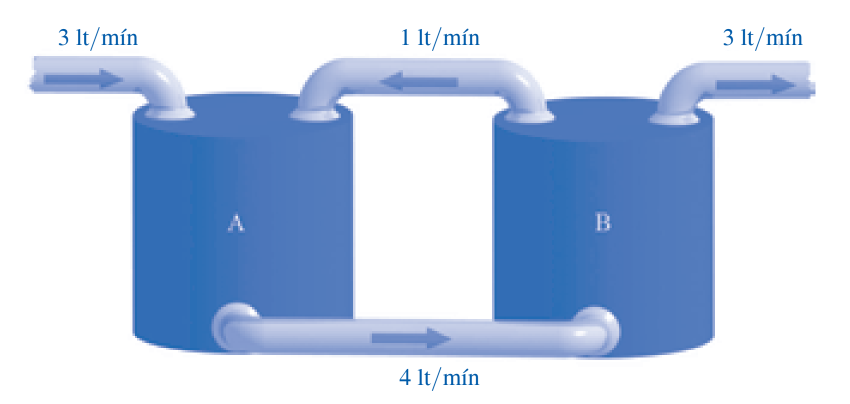


Figura 6-20 •

Supongamos que se tienen dos depósitos inicialmente llenos de agua con 40 litros cada uno, el tanque A contiene 20 kilogramos de sal disuelta. En $t = 0$ comienza a fluir agua pura al tanque A a razón de 3 litros por minuto; simultáneamente, por medio de bombas, fluye agua del tanque A hacia el B, del tanque B al A y del tanque B hacia afuera con los ritmos que se muestran en la figura 6-20. El objetivo de este problema es determinar las cantidades x_1 y x_2 de sal que se encuentran en sendos depósitos en cada instante del tiempo t; se supone que en todo momento las mezclas permanecen homogéneas en los dos recipientes. Plantear el problema no es tan complicado; pero resolverlo tiene algunas dificultades que, afortunadamente, con álgebra lineal se pueden soslayar. Sean r_1 y r_2 los ritmos a los que entra y sale sal del tanque A, respectivamente. Entonces,

$$\frac{dx_1}{dt} = r_1 - r_2.$$

Note que en cada tanque, debido a los flujos que se muestran en la figura 6-20, los volúmenes permanecen constantes en la medida que transcurre el tiempo. Así, del tanque B al tanque A fluyen $1(x_2/40)$ kilogramos de sal por minuto. Puesto que del tanque A sale agua a razón de 4 litros por minuto, entonces del tanque A escapa sal (hacia el tanque B) a razón de $4(x_1/40)$ kilogramos de sal por minuto. Por tanto,

$$r_1 = \frac{1}{40}x_2$$

$$r_2 = \frac{4}{40}x_1$$

entonces

$$\frac{dx_1}{dt} = -\frac{1}{10}x_1 + \frac{1}{40}x_2.$$

De manera análoga

$$\frac{dx_2}{dt} = \rho_1 - \rho_2$$

donde ρ_1 es la rapidez con la que entra sal al recipiente B y ρ_2 es con la que escapa sal del recipiente B. Por lo que acabamos de explicar, $\rho_1 = 4(x_1/40)$, y $\rho_2 = 1(x_2/40) + 3(x_2/40)$; por tanto,

$$\frac{dx_2}{dt} = \frac{1}{5}x_1 - \frac{1}{40}x_2 - \frac{3}{40}x_2$$

$$= \frac{1}{10}x_1 - \frac{1}{10}x_2$$

De esta manera el problema consiste en resolver el sistema de ecuaciones diferenciales

$$\frac{dx_1}{dt} = -\frac{1}{10}x_1 + \frac{1}{40}x_2$$

$$\frac{dx_2}{dt} = \frac{1}{10}x_1 - \frac{1}{10}x_2$$

con la condición inicial $x_1(0) = 20$, $x_2(0) = 0$.

Tratemos el caso en general. Un sistema lineal de n ecuaciones con n incógnitas tiene la forma

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ &\vdots \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{aligned} \tag{6.57}$$

Donde las x_i son funciones incógnitas, $\dot{x}_i$ representa la derivada de x_i respecto a t; es decir, $\dot{x}_i = \dfrac{dx_i}{dt}$; y los a_{ij} son números reales dados. Es claro que el sistema lineal de ecuaciones diferenciales se puede escribir matricialmente como

$$\dot{\vec{x}} = A\vec{x} \tag{6.58}$$

con:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \dot{\vec{x}} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}.$$

1. Por el momento, supongamos que tenemos únicamente una ecuación diferencial con una incógnita; entonces el sistema (6.57) se reduce a

$$\dot{y} = \lambda y$$

Esto es

$$\frac{dy}{dt} = \lambda y.$$

Entonces,

$$\begin{aligned} \frac{dy}{dt} = \lambda y \quad &\Rightarrow \quad \frac{dy}{y} = \lambda dt \\ &\Rightarrow \quad \int \frac{dy}{y} = \lambda \int dt \\ &\Rightarrow \quad \ln|y| = \lambda t + C \\ &\Rightarrow \quad y = C_1 e^{\lambda t}. \end{aligned}$$

La solución es $y = C_1 e^{\lambda t}$ para alguna constante C_1.

2. Supongamos que A es diagonal, esto es $A = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Así, el sistema (6.58) se reduce a

$$\begin{aligned} \dot{x}_1 &= \lambda_1 x_1 \\ \dot{x}_2 &= \lambda_2 x_2 \\ &\vdots \\ \dot{x}_n &= \lambda_n x_n \end{aligned}$$

y, por tanto, del inciso anterior, $x_i = C_i e^{\lambda_i t}$, donde C_i es una constante.

3. Supongamos finalmente que A es diagonalizable con $\lambda_1, \lambda_2, \ldots, \lambda_n$ sus valores propios y M la correspondiente diagonalización:

$$M^{-1}AM = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} = D.$$

De esta manera,

$$A = MDM^{-1},$$

luego

$$\begin{aligned} \dot{\vec{x}} &= A\vec{x} \\ &= (MDM^{-1})\vec{x} \\ &= (MD)(M^{-1}\vec{x}), \end{aligned}$$

y, por tanto,

$$M^{-1}\dot{\vec{x}} = D(M^{-1}\vec{x});$$

de donde, si $\vec{y} = M^{-1}\vec{x}$, entonces $\dot{\vec{y}} = M^{-1}\dot{\vec{x}}$ y el sistema se transforma a

$$\dot{\vec{y}} = D\vec{y}$$

el cual tiene, por el segundo inciso, solución

$$\vec{y} = \begin{bmatrix} C_1 e^{\lambda_1 t} \\ C_2 e^{\lambda_2 t} \\ \vdots \\ C_n e^{\lambda_n t} \end{bmatrix},$$

y así,

$$\vec{x} = M \begin{bmatrix} C_1 e^{\lambda_1 t} \\ C_2 e^{\lambda_2 t} \\ \vdots \\ C_n e^{\lambda_n t} \end{bmatrix}.$$

es la solución del sistema (6.57) si la matriz A de (6.58) es diagonalizable.

► **Ejemplo 6.47** Resolver el problema de los tanques y mezclas enunciado al inicio de esta sección.◄

Solución Recordemos que el sistema para este modelo es

$$\frac{dx_1}{dt} = -\frac{1}{10}x_1 + \frac{1}{40}x_2$$
$$\frac{dx_2}{dt} = \frac{1}{10}x_1 - \frac{1}{10}x_2$$

que en forma matricial se escribe

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1/10 & 1/40 \\ 1/10 & -1/10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

La matriz de coeficientes del sistema es

$$A = \begin{bmatrix} -1/10 & 1/40 \\ 1/10 & -1/10 \end{bmatrix}.$$

El polinomio caractecterístico para esta matriz es

$$p_A(\lambda) = \lambda^2 + \frac{1}{5}\lambda + (\frac{1}{100} - \frac{1}{400})$$
$$= \lambda^2 + \frac{1}{5}\lambda + \frac{3}{400}$$

cuyas raíces son $\lambda_1 = -1/20$, $\lambda_2 = -3/20$; por tanto (ya que los valores propios son distintos) A es diagonalizable.

- Para $\lambda_1 = -1/20$,

$$\begin{aligned}[A-\lambda_1 I_2] &= \begin{bmatrix} -1/10+1/20 & 1/40 \\ 1/10 & -1/10+1/20 \end{bmatrix} \\ &= \begin{bmatrix} -1/20 & 1/40 \\ 1/10 & -1/20 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & -1/2 \\ 1/10 & -1/20 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & -1/2 \\ 0 & 0 \end{bmatrix}\end{aligned}$$

y, por tanto,

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = r \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$$

son, con $r \neq 0$, los vectores propios correspondientes a $\lambda_1 = -1/20$.

- Si $\lambda_2 = -3/20$, entonces

$$\begin{aligned}[A-\lambda_2 I_2] &= \begin{bmatrix} -1/10+3/20 & 1/40 \\ 1/10 & -1/10+3/20 \end{bmatrix} \\ &= \begin{bmatrix} 1/20 & 1/40 \\ 1/10 & 1/20 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1/2 \\ 1/10 & 1/20 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1/2 \\ 0 & 0 \end{bmatrix}\end{aligned}$$

Por tanto, los vectores propios correspondientes a $\lambda_2 = -3/20$, tienen la forma

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = r \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}, \; r \neq 0.$$

Así, una diagonalización para A está dada por el par $(M, \text{diag}(-1/20, -3/20))$, donde

$$M = \begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix}$$

Luego,

$$\begin{aligned}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= M \begin{bmatrix} C_1 e^{-t/20} \\ C_2 e^{-3t/20} \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} C_1 e^{-t/20} \\ C_2 e^{-3t/20} \end{bmatrix} \\ &= \begin{bmatrix} C_1 e^{-t/20} - C_2 e^{-3t/20} \\ 2C_1 e^{-t/20} + 2C_2 e^{-3t/20} \end{bmatrix}\end{aligned}$$

Esto es,

$$\begin{aligned} x_1 &= C_1 e^{-t/20} - C_2 e^{-3t/20} \\ x_2 &= 2C_1 e^{-t/20} + 2C_2 e^{-3t/20} \end{aligned}$$

Finalmente, se debe cumplir $x_1(0) = 20$ y $x_2(0) = 0$, por tanto,

$$C_1 - C_2 = 20$$
$$2C_1 + 2C_2 = 0$$

que al resolver produce $C_1 = 10$ y $C_2 = -10$. La solución está dada por

$$\begin{aligned} x_1 &= 10(e^{-t/20} + e^{-3t/20}) \\ x_2 &= 20(e^{-t/20} - e^{-3t/20}). \end{aligned}$$ ✔

► **Ejemplo 6.48** Resolver el sistema lineal de ecuaciones diferenciales

$$\begin{aligned} \dot{x}_1 &= x_1 - 3x_2 + 3x_3 \\ \dot{x}_2 &= -5x_2 + 6x_3 \\ \dot{x}_3 &= -3x_2 + 4x_3 \end{aligned}$$ ◄

Solución En este caso la matriz de coeficientes es

$$A = \begin{bmatrix} 1 & -3 & 3 \\ 0 & -5 & 6 \\ 0 & -3 & 4 \end{bmatrix}$$

Por el ejemplo 5.50, pág. 478, A es diagonalizable con:

$$M = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

Entonces,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} C_1 e^t \\ C_2 e^t \\ C_3 e^{-2t} \end{bmatrix} = \begin{bmatrix} C_1 e^t + C_3 e^{-2t} \\ C_2 e^t + 2C_3 e^{-2t} \\ C_2 e^t + C_3 e^{-2t} \end{bmatrix}$$ ✔

6.8 Optimización de funcionales

Requisitos: *Conocimientos de cálculo diferencial y nociones de cálculo integral de una variable; espacios vectoriales normados y transformaciones lineales: 4.2.1, 4.2.2, 4.2.4 y 5.1.1.*

Existe una gran variedad de problemas de aplicación en los que la solución consiste en calcular el valor numérico máximo o mínimo que tiene un proceso, se llaman problemas de optimización. Ya hemos estudiado un tipo especial de ellos en la sección de programación lineal. Aquí analizaremos la optimización de funcionales; esto es, de funciones cuyos argumentos pertenecen a un espacio vectorial y sus valores son números reales. Seguramente el lector está familiarizado con problemas de una variable real en los que se tienen que encontrar los máximos y mínimos relativos de una función. La herramienta que utilizó para resolverlos fue la derivación, calculando los puntos críticos –donde la derivada vale cero–

y empleando algún criterio más para saber si en ellos se alcanza un máximo o un mínimo. En el caso general tendremos un funcional J con dominio contenido en un espacio vectorial normado y valores reales; entonces para encontrar los valores máximos o mínimos del funcional, necesitaremos generalizar el concepto de derivación de funciones de una variable real a este tipo de funcionales. Aquí veremos la utilidad que tienen los espacios vectoriales normados, en particular los espacios $C^n[a,b]$ que estudiamos en el capítulo 4; también veremos la importancia que tienen los conceptos de transformación lineal y valores propios que vimos en el capítulo 5. Trabajaremos en un contexto general y veremos con detalle el caso de la teoría de diferenciación en los espacios $\mathbb{R}^k$ y la optimización de funcionales con dominio en estos espacios mediante los valores propios de la matriz hessiana. Sin embargo, también trataremos con la diferenciación y optimización de funcionales en espacios vectoriales de dimensión infinita. Antes, en el siguiente apartado, daremos varios ejemplos de problemas físicos donde se presenta la necesidad de optimizar un funcional para poder resolverlos.

6.8.1 *Problemas físicos*

Principio de Fermat y ley de Snell

El principio de Fermat establece que la trayectoria que seguirá un haz de luz desde un punto dado en un medio hasta otro punto en otro medio, será aquella que ocupe el tiempo mínimo. En la figura 6-21, un rayo de luz se mueve desde un punto A en un medio con una velocidad v_1, pasando por el punto P y, a partir de ahí, se mueve en otro medio más denso con velocidad v_2 hasta llegar a un punto B. De esta figura se infiere que el tiempo total empleado para llegar desde el punto A al punto B es

$$T(x) = \frac{\sqrt{a^2 + x^2}}{v_1} + \frac{\sqrt{(c-x)^2 + b^2}}{v_2}.$$

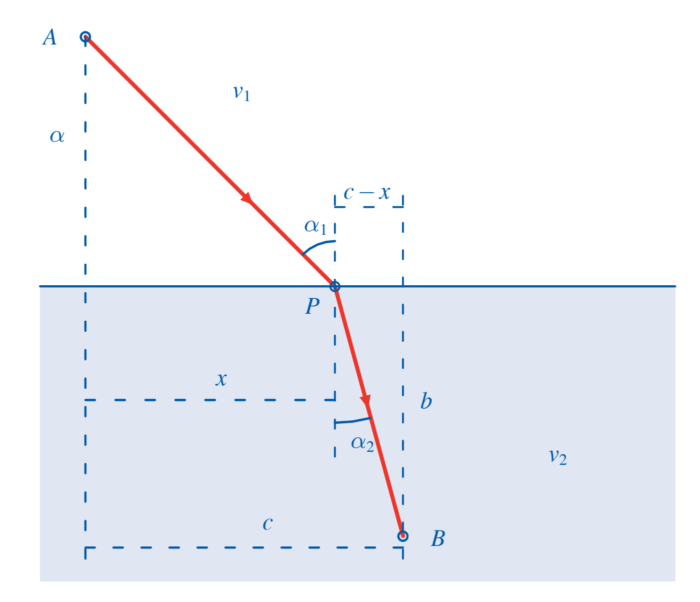


Figura 6-21 • Rayo de luz refractándose al pasar de un medio a otro.

Si aceptamos el principio de Fermat y se alcanza un mínimo en x, entonces,

$$T'(x) = 0$$

lo cual implica

$$\frac{x}{v_1\sqrt{a^2+x^2}} = \frac{c-x}{\sqrt{(c-x)^2+b^2}};$$

esto es,

$$\frac{\operatorname{sen}\alpha_1}{v_1} = \frac{\operatorname{sen}\alpha_2}{v_2}.$$

Esta es la ley de refracción de Snell y es un problema que se resuelve basándose en hallar las relaciones necesarias entre los parámetros que aparecen en el planteamiento, a partir de hallar puntos críticos de una función.

Braquistócrona y la solución de John Bernoulli

Quizás el primer problema que da origen al cálculo de variaciones es el de la braquistócrona, que fue planteado por John Bernoulli como un desafío a finales del siglo XVII. El problema consiste en encontrar la curva de descenso más rápido entre dos puntos dados en el plano. En la figura 6-22 se encuentran dos puntos en el plano, A y B, y una partícula desciende a través de una curva desde el punto A hasta el punto B. Se supone que la única fuerza que actúa es la gravedad. El desafío consiste en encontrar, de entre todas las curvas que pasan por A y B, aquella para la cual el tiempo de descenso es el mínimo. Newton fue el primero en responder "correctamente" este desafío; la solución del propio John al problema la presentamos a continuación.

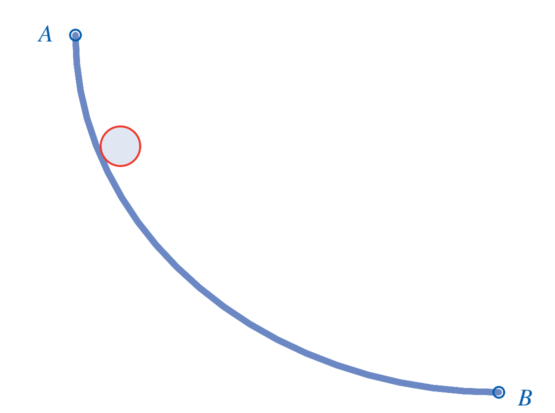


Figura 6-22 • El problema de la braquistócrona.

Regresemos a la ley de refracción de Snell, ahora supongamos, además, que en lugar de tener solamente un medio, tenemos varios medios donde la luz se mueve con velocidades distintas, pero constantes, en cada uno de ellos, como se indica en la figura 6-23. Entonces tendremos la relación

$$\frac{\operatorname{sen}\alpha_1}{v_1} = \frac{\operatorname{sen}\alpha_2}{v_2} = \frac{\operatorname{sen}\alpha_3}{v_3} = \frac{\operatorname{sen}\alpha_4}{v_4}$$

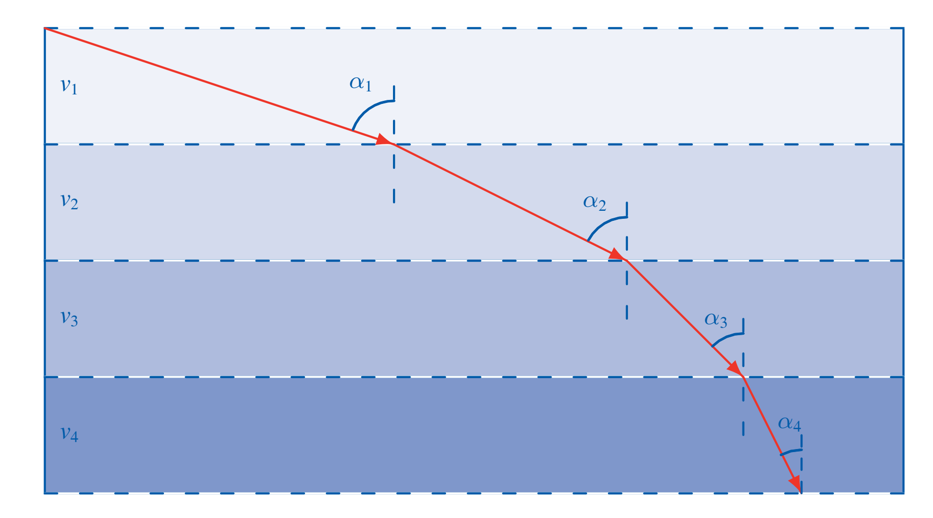


Figura 6-23 • Rayo de luz refractándose al pasar varios medios.

Ahora, si los medios son cada vez más numerosos y más delgados, entonces en el límite la velocidad de la luz decrece continuamente mientras el rayo desciende y se concluye

$$\frac{\operatorname{sen}\alpha}{v} = c$$

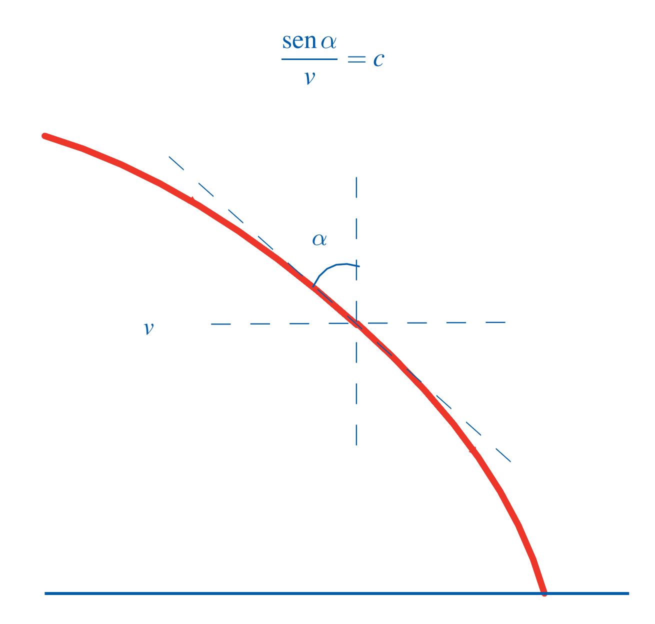


Figura 6-24 • Trayectoria curva de la luz.

como se ilustra en la figura 6-24. Esto sucede en la práctica, aproximadamente, cuando la luz desciende hacia la Tierra a través de una atmósfera de densidad creciente. Regresando al problema de Bernoulli, pensemos en un sistema de referencia con dirección vertical positiva orientada hacia abajo y el punto A situado en el origen, como se ilustra en la figura 6-25. John Bernoulli imaginó que, así como para un rayo de luz la naturaleza "selecciona" la trayectoria que se recorre en el tiempo más corto, también "selecciona" la trayectoria para que una partícula se deslice de A a B en el menor tiempo. Si esto es así, se debe tener nuevamente

$$\frac{\operatorname{sen}\alpha}{v} = c \tag{6.59}$$

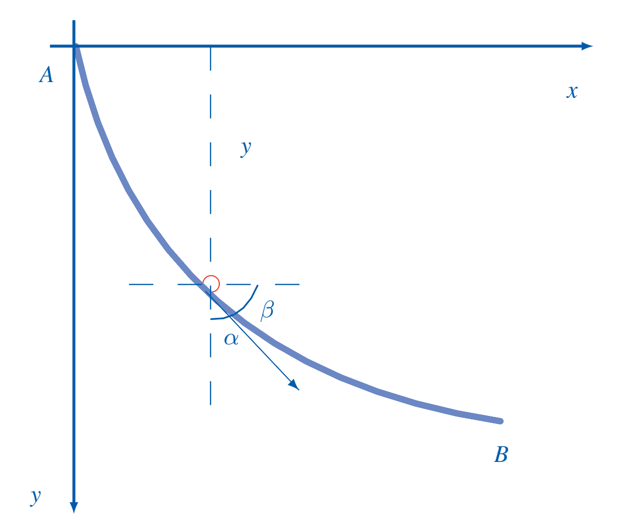


Figura 6-25 • La braquistócrona.

Como que se supone que no hay otras fuerzas que influyan en el movimiento de la partícula, hay conservación de la energía y el trabajo realizado por la fuerza gravitacional para llevar la partícula del punto A al punto B es la variación de la energía cinética. Entonces, si v representa el módulo de la velocidad, m la masa de la partícula, g la aceleración debida a la gravedad, y el desplazamiento vertical y W el trabajo,

$$mgy = W = \frac{1}{2}mv^2$$

y de esta última relación se desprende

$$v = \sqrt{2gy} \tag{6.60}$$

De la figura 6-25 se puede ver que

$$\begin{aligned}\operatorname{sen}^2\alpha &= \cos^2\beta \\ &= \left(1+\tan^2\beta\right)^{-1} \\ &= \left(1+(y')^2\right)^{-1}\end{aligned}$$

y de (6.59), (6.60) y la precedente igualdad se deduce

$$y\left(1+(y')^2\right) = C \tag{6.61}$$

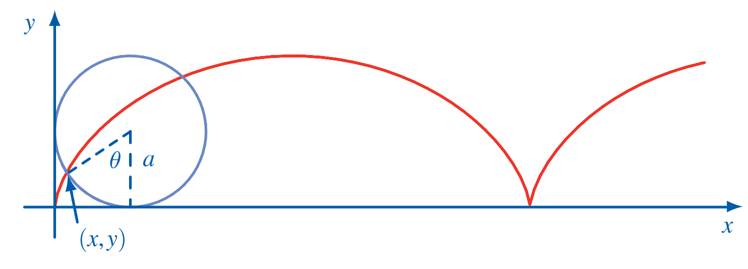


Figura 6-26 • Cicloide, curva que describe la trayectoria de un punto dado de la circunferencia de radio a cuando ésta rueda sobre el eje x. Es la curva solución para el problema de la braquistócrona según Newton y Bernoulli.

donde C es una constante. Entonces, la curva que resuelve el problema, de acuerdo con Bernoulli, debe ser la solución de la ecuación diferencial ordinaria no lineal (6.61). A esta ecuación diferencial le llamaremos la ecuación de la braquistócrona. Al resolver esta ecuación se obtiene

$$\begin{aligned} x &= a(\cos\theta - \operatorname{sen}\theta) \\ y &= a(1-\cos\theta) \end{aligned}$$

en forma paramétrica, cuyo lugar geométrico es la cicloide generada por la circunferencia de radio a, como se muestra en la figura 6-26.

Enfoque "moderno" al problema de la braquistócrona

Pensemos nuevamente en el problema planteado por J. Bernoulli, pero enfocándolo en forma distinta. Supongamos que la curva mostrada en la figura 6-25 es la gráfica de una función $y = f(x)$ que pasa por los puntos $A = (0,0)$ y $B = (b, y_1)$. De nuevo, por el principio de conservación de la energía, $v = \sqrt{2gy}$. Pero si $s(t)$ es la longitud de arco para $y = f(x)$, entonces $v = \dfrac{ds}{dt}$; luego

$$\begin{aligned} \frac{ds}{dx}\frac{dx}{dt} &= v \\ &= \sqrt{2gy} \end{aligned}$$

esto es (si suponemos que $y = f(x)$ es derivable y continua en $[0,b]$)

$$\sqrt{1+[y']^2}\frac{dx}{dt} = \sqrt{2gy}.$$

De donde el tiempo total para desplazarse del punto A al punto B, a través de la curva $y = f(x)$, está dado por

$$J(y) = \frac{1}{\sqrt{2g}}\int_0^b \sqrt{\frac{1+[y']^2}{y}}dx.$$

Hemos establecido así un funcional J, que a cada función continua y en $[0,b]$, con la condición $y(0) = 0$ y $y(b) = y_1$, le asigna el tiempo de recorrido por la partícula al descender a través de la curva $y = f(x)$. Tenemos entonces que hallar, de entre todos los elementos del dominio de J, aquella función para la que J toma un valor mínimo. Este es el tipo de problemas que analizaremos en este apartado por medio de la generalización del concepto de derivada de una función de una variable a diferencial de funcionales.

Problema isoperimétrico

Los antiguos griegos plantearon el problema de encontrar, de entre todas las curvas planas simples y cerradas de longitud dada, aquella que encierra la mayor área. En forma más o menos rigurosa llegaron a la conclusión que dicha curva es el círculo. Supongamos que C es una curva plana simple y cerrada con parametrización $(x(t), y(t))$, $t \in [a,b]$. Pongamos $F(x,y) = (P(x,y), Q(x,y)) = (-y, x)$ y apliquemos el teorema de Green en el plano:

$$\int_C Pdx + Qdy = \iint_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy$$

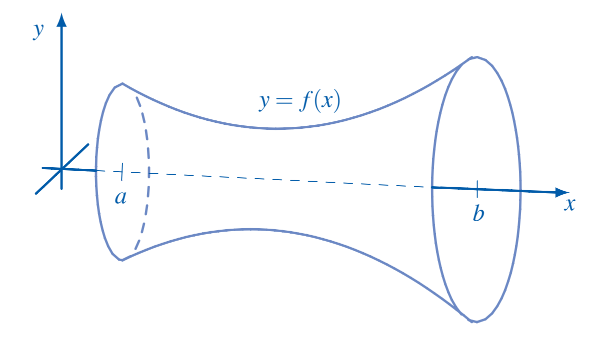


Figura 6-27 • Superficie de revolución generada por la gráfica de $y = f(x)$.

donde A es la región encerrada por la curva C. Entonces,

$$\int_a^b (-y,x)\cdot(x',y') = 2\iint_A dxdy\,.$$

Luego, si representamos por $|A|$ el área de la región A, se tiene

$$|A| = \frac{1}{2}\int_a^b (xy' - x'y)\,dt\,.$$

Se tienen que buscar un par de funciones derivables y con continuidad x,y en el intervalo $[a,b]$, tales que la integral en el lado derecho de la igualdad es máxima y, además, que estas funciones cumplan

$$\int_a^b \sqrt{[x']^2 + [y']^2}dt = L$$

para L dada. Note que nuevamente hemos establecido un problema de optimizar un funcional, en este caso encontrar un valor máximo.

Superficie de revolución de área mínima

En este problema se trata de encontrar, de entre todas las funciones $y = f(x)$ cuyas gráficas pasan por los puntos $A = (a,y_1)$, $B = (b,y_2)$, aquélla para la cual la superficie de revolución que se obtiene al girar la gráfica de f en $[a,b]$ alrededor del eje de las abscisas tiene área mínima (vea la figura 6-27). De cálculo sabemos que el área de la superficie del sólido de revolución está dada por

$$S(f) = 2\pi\int_a^b f(x)\left(1 + [f'(x)]^2\right)^{1/2} dx\,.$$

Aquí otra vez el problema se reduce a encontrar una función que minimice el precedente funcional, para aquellas funciones f que sean derivables con continuidad en $[a,b]$ y que satisfagan la condición $f(a) = y_1$, $f(b) = y_2$.

Geodésicas en el plano

Dados dos puntos en el plano (a,y_1), (b,y_2), encontrar la curva de longitud mínima que los una. En este caso el funcional a minimizar es

$$L(f) = \int_a^b \sqrt{1 + [f'(x)]^2}dx\,.$$

Geodésicas en una superficie en el espacio

Dados dos puntos M_0 y M_1 en una superficie, se quiere encontrar la curva (geodésica) en la superficie que tenga la mínima longitud y que los una (vea la figura 6-28). Si una superficie está dada en forma paramétrica por una función $M(u,v)$ con (u,v) en una región del plano, entonces toda curva $M(t)$ sobre la superficie proviene de una curva $\gamma(t) = (u(t), v(t))$, $t \in [0,1]$, cuya longitud es entonces

$$\int_0^1 \left\| \frac{dM}{dt} \right\| dt\,.$$

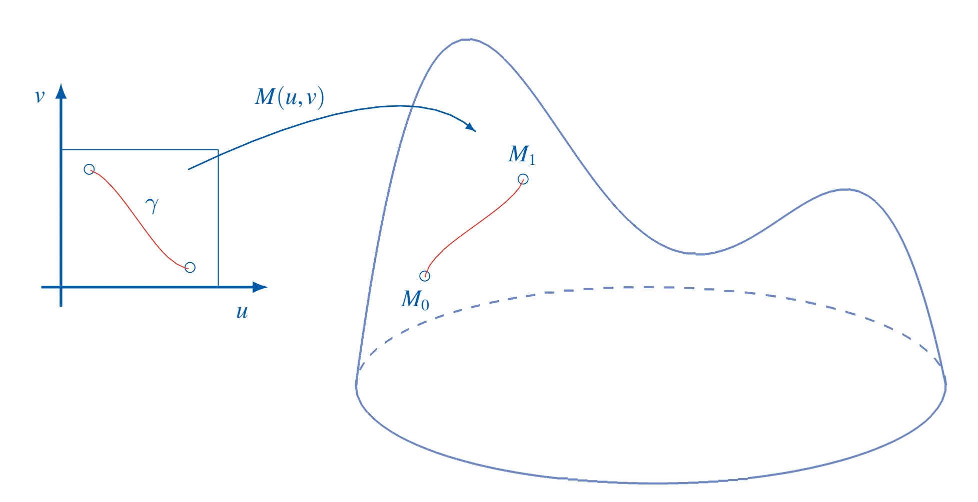


Figura 6-28 • Toda curva sobre la superficie proviene de una curva sobre el plano. Una geodésica es una curva de longitud mínima en una superficie que une a dos puntos y sirve para definir distancia entre puntos de la superficie.

El funcional a minimizar en este caso es

$$L(u,v) = \int_0^1 \|M_u u' + M_v v'\|\, dt$$

con la condición $M(0) = M_0$ y $M(1) = M_1$; donde M_u y M_v son las derivadas parciales de M.

Catenarias

Los extremos de un cable se suspenden de dos puntos a la misma altura y se deja colgar por su propio peso. ¿Qué forma adquiere? Sea $y = f(x)$ la función cuya gráfica coincide con la forma que tiene el cable en las condiciones descritas. Vamos a suponer que la densidad por unidad de longitud del cable es constante en todo punto con magnitud ρ. Entonces, si ds es un infinitesimal de longitud del cable, éste contribuye con un infinitesimal de energía potencial

$$dW = \rho g y ds\,,$$

donde hemos tomado como sistema de referencia el sistema de ejes cartesianos de la figura 6-29 y g es la constante de aceleración gravitacional. Entonces, la energía potencial total está dada por

$$W = \int_{-a}^{a} \rho g y ds = \rho g \int y\sqrt{1+[y']^2}dx.$$

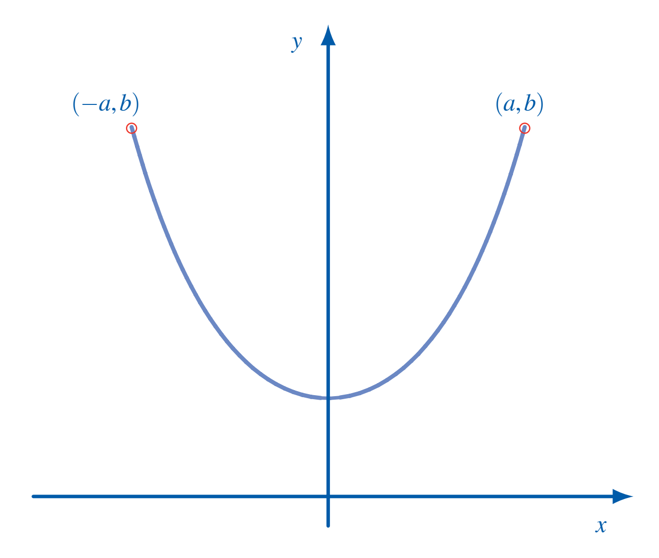


Figura 6-29 • Catenaria, forma que toma un cable cuando se deja colgar fijando sus extremos a dos puntos.

Si consideramos nuevamente que la naturaleza mínimiza recursos, la forma que debe adquirir el cable es aquella para la cual se tenga una energía potencial mínima. Por tanto, se debe encontrar el mínimo del funcional

$$J(y) = \rho g \int y\sqrt{1+[y']^2}dx$$

donde $y \in C^1[a,b]$, $y(-a) = y(a) = b$. A la curva solución de este problema se le llama **catenaria**.

6.8.2 Cálculo diferencial en espacios vectoriales

En esta sección vamos a generalizar el concepto de derivada de una función de variable real, no sólo a funciones de varias variables, sino a funciones definidas en conjuntos contenidos en espacios vectoriales, aun si éstos tienen dimensión infinita. Una de las características más importantes de este tema es que analizaremos transformaciones sobre espacios de funciones; esto es, funciones cuyos argumentos son funciones. Por tanto, no usaremos la notación de poner flechas encima de los vectores para distinguirlos de los escalares y simplemente utilizaremos letras minúsculas para representar a los elementos de los espacios en los que tengamos que trabajar y mayúsculas para las funciones que tengan por argumentos a dichos elementos. Regresaremos a la notación usual cuando sea necesario en lo casos particulares en los que aparecen los espacios $\mathbb{R}^k$.

Límites y continuidad en espacios normados

En lo que sigue **E** y **F** son un par de espacios vectoriales normados reales y las normas para cada uno de ellos se representan con el mismo símbolo, $\|\cdot\|$, a menos que se indique otra cosa. Además, a los neutros aditivos de ambos espacios y al escalar nulo del campo $\mathbb{R}$, los denotaremos a todos por el símbolo 0; esta es una medida necesaria de comodidad y hasta claridad a estas alturas; en el contexto siempre quedará claro de qué neutro aditivo se trata.

Definición 6.8 *Sean J una función con dominio contenido en $\mathbf{E}$ y valores en $\mathbf{F}$, $f_0 \in \mathbf{E}$ tal que f está definida en todos los puntos de una bola abierta $B(f_0, r)$ excepto, quizá, en el propio centro f_0. Se dice que la función J tiene como límite a $l \in \mathbf{F}$ cuando f tiende a f_0, lo cual se denota por*

$$\lim_{f \to f_0} J(f) = l \,,$$

si para toda bola abierta[14] *$B(l, \varepsilon)$, existe una bola abierta $B(f_0, \delta)$ contenida en $B(f_0, r)$ tal que*

$$J(f) \in B(l, \varepsilon)\; \forall f \in B(f_0, \delta) - \{f_0\}$$

Esto es,

$$0 < \|f - f_0\| < \delta \Rightarrow \|J(f) - l\| < \varepsilon \,.$$

Se puede probar que una función tiene a lo más un límite cuando sus argumentos tienden a un elemento dado y que, además, si se cambian las normas en $\mathbf{E}$ y $\mathbf{F}$ por sendas normas equivalentes,[15] entonces $\lim_{f \to f_0} J(f) = l$ respecto a las normas originales si y sólo si $\lim_{f \to f_0} J(f) = l$ respecto a las normas equivalentes.

Note que

$$\lim_{f \to f_0} J(f) = l \Leftrightarrow \lim_{f \to f_0} \|J(f) - l\| = 0.$$

▶ **Ejemplo 6.49** Claramente:

1. $\lim_{f \to f_0} g = g$.
2. $\lim_{f \to f_0} f = f_0$. ◀

▶ **Ejemplo 6.50** Sea $J : \mathbb{R} \to \mathbf{E}$, el operador definido por

$$J(t) = f + th$$

donde $f, h \in \mathbf{E}$ son vectores fijos. Entonces, puesto que

$$\begin{aligned} \lim_{t \to 0} \|f + th - f\| &= \lim_{t \to 0} \|th\| \\ &= \lim_{t \to 0} |t| \|h\| \\ &= 0 \,, \end{aligned}$$

$$\lim_{t \to 0} J(t) = f \,;$$

es decir,

$$\lim_{t \to 0} (f + th) = f \,. ◀$$

[14] Recuerde que $B(f_0, r) = \{f \mid \|f - f_0\| < r\}$ (cfr. definición 4.13, pág.311).
[15] Cfr. definición 4.14, pág. 324.

Definición 6.9 *Sean $D \subset \mathbf{E}$, $J : D \to \mathbf{F}$ una función y $f_0 \in D$. La función J es* ***continua*** *en f_0 si*

$$\lim_{f \to f_0} J(f) = J(f_0).$$

Observe que J es continua en f_0 si y sólo si $\lim_{f \to f_0} J(f_0 + h) = J(f_0)$. Nuevamente, si se cambian las normas de cada espacio por normas equivalentes, la función es continua en un punto respecto a las normas originales si y sólo si es continua respecto a las normas equivalentes.

❍ **Nota 6.4** No es difícil mostrar que en $C^n[a,b]$, el espacio de funciones con derivadas hasta el orden n continuas en el intervalo $[a,b]$,

$$\|f\|_n = \max\{\|f\|_\infty, \|f'\|_\infty, \ldots, \|f^{(n)}\|_\infty\}$$

es una norma (cfr. ejercicio 312 del capítulo 4). A lo largo de toda esta sección estaremos utilizando los espacios $C^n[a,b]$ y, a menos que se indique lo contrario, la norma que utilizaremos será siempre ésta. Cuando $n = 0$, nos referimos simplemente al espacio $C[a,b]$ con la norma uniforme $\|\cdot\|_0 = \|\cdot\|_\infty$.

► **Ejemplo 6.51** Sea $J : C^1[0,1] \to \mathbb{R}$ el funcional definido por

$$J(y) = \int_0^1 (y + 2y')\, dx$$

demostrar que J es continuo en f_0, donde $f_0(x) = x$ para todo $x \in [a,b]$.◄

DEMOSTRACIÓN ■ Sea $y = y(x) \in C^1[0,1]$ y $\varepsilon > 0$ dado, se tiene

$$\begin{aligned}
|J(y) - J(f_0)| &= \left| \int_0^1 (y + 2y')\, dx - \int_0^1 (x + 2)\, dx \right| \\
&= \left| \int_0^1 (y - x)\, dx + \int_0^1 (2y' - 2)\, dx \right| \\
&\le \left| \int_0^1 (y - x)\, dx \right| + \left| \int_0^1 (2y' - 2)\, dx \right| \\
&\le \int_0^1 |y - x|\, dx + 2 \int_0^1 |(y' - 1)|\, dx.
\end{aligned}$$

Entonces, si $\|y - f_0\| < \varepsilon/3$,

$$\int_0^1 |y - x|\, dx < \int_0^1 \frac{\varepsilon}{3} dx = \frac{\varepsilon}{3},$$

$$2 \int_0^1 |(y' - 1)|\, dx < 2 \int_0^1 \frac{\varepsilon}{3} dx = \frac{2}{3}\varepsilon$$

y, por tanto,

$$\begin{aligned}
|J(y) - J(f_0)| &\le \int_0^1 |y - x|\, dx + 2 \int_0^1 |(y' - 1)|\, dx \\
&< \frac{\varepsilon}{3} + \frac{2}{3}\varepsilon \\
&= \varepsilon.
\end{aligned}$$

Así, con $\delta = \varepsilon/3$,

$$\|y - f_0\| < \delta \Rightarrow |J(y) - J(f_0)| < \varepsilon,$$

lo cual significa

$$\lim_{f_0 \to f} J(f) = J(f_0)$$

y por tanto la función J es continua en f_0. ■

De manera análoga a las funciones reales de una variable real se cumplen las propiedades enunciadas en el siguiente teorema.

Teorema 6.7 *Sean* **E**, **F** *y* **G** *espacios vectoriales normados.*

1. *Sean* $J_1, J_2 : D \subset \mathbf{E} \to \mathbf{F}$ *un par de funciones y* $k \in \mathbb{R}$. *Si* $\lim_{f \to f_0} J_1(f) = l_1$ *y* $\lim_{f \to f_0} J_2(f) = l_2$, *entonces:*

   (a) $$\lim_{f \to f_0} (J_1(f) \pm J_2(f)) = l_1 \pm l_2.$$

   (b) $$\lim_{f \to f_0} kJ_1(f) = kl_1.$$

   (c) *Si*[16] $\mathbf{F} = \mathbb{R}$,

   (i) $$\lim_{f \to f_0} J_1(f)J_2(f) = l_1 l_2.$$

   (ii) $$\lim_{f \to f_0} \frac{J_1(f)}{J_2(f)} = \frac{l_1}{l_2}.$$

   si $l_2 \neq 0$

   (d) *Si* $\mathbf{F} = C^n[a,b]$,

   $$\lim_{f \to f_0} J_1(f)J_2(f) = l_1 l_2.$$

2. *Sean* $J_2 : D_2 \subset \mathbf{F} \to \mathbf{G}$ *y* $J_1 : D_1 \subset \mathbf{E} \to D_2$. *Si* $\lim_{f \to f_0} J_1(f) = l$ *y* J_2 *es continua en* l, *entonces*

   $$\lim_{f \to f_0} J_2(J_1(f)) = J_2(l).$$

Como consecuencia inmediata tenemos el siguiente teorema.

[16] Las funciones J_1J_2 y J_1/J_2 se definen como $(J_1J_2)(f) = J_1(f)J_2(f)$ y $(J_1/J_2)(f) = J_1(f)/J_2(f)$, respectivamente.

Teorema 6.8 *Sean* $\mathbf{E}$, $\mathbf{F}$ *y* $\mathbf{G}$ *espacios vectoriales normados.*

1. *Sean* J_1, J_2 *funciones definidas en un subconjunto de* $\mathbf{E}$ *con valores en* $\mathbf{F}$ *que son continuas en* f_0 *y* k *un escalar:*

   (a) $J_1 \pm J_2$ *es continua en* f_0.

   (b) kJ_1 *es continua en* f_0.

   (c) $J_1 J_2$ *es continua en* f_0*, si* $\mathbf{F} = \mathbb{R}$ *o* $\mathbf{F} = C^n[a,b]$.

   (d) J_1/J_2 *es continua en* f_0*, si* $\mathbf{F} = \mathbb{R}$ *y* $J_2(f_0) \neq 0$.

2. *Sean* $J_2 : D_2 \subset \mathbf{F} \to \mathbf{G}$ *y* $J_1 : D_1 \subset \mathbf{E} \to D_2$. *Si* J_1 *es continua en* f_0 *y* J_2 *es continua en* $J_1(f_0)$*, entonces la función composición* $f \mapsto J_2(J_1(f))$ *es continua en* f_0.

El teorema precedente afirma que la continuidad es una característica que se preserva bajo las operaciones usuales de funciones; es decir, la suma, diferencia, producto de funciones, cociente de funciones y composición de funciones continuas también dan como resultado funciones continuas, siempre que dichas operaciones estén definidas.

► **Ejemplo 6.52** Sea la función $J : C^1[a,b] \to C[a,b]$ definida por

$$J(f) = 1 + f'.$$

Mostremos que J es una función continua en cualquier elemento del espacio $C^1[a,b]$ (recuerde que la norma en este espacio está definida por $\|f\| = \max\{\|f\|_\infty, \|f'\|_\infty\}$; mientras que en $C[a,b]$ se trabaja con la norma uniforme $\|\cdot\|_\infty$). Sea la función $J_1 : C^1[a,b] \to C[a,b]$ definida por $J_1(f) = f'$; afirmamos que J_1 es continua en todo punto $f_0 \in C^1[a,b]$. Efectivamente

$$\|J_1(f) - J_1(f_0)\|_\infty = \|f - f'\|_\infty \leq \|f - f_0\|.$$

y, por tanto, $\lim_{f\to f_0}(J_1(f) - J_1(f_0)) = 0$. Claramente toda función constante es continua en todo punto y como $J = 1 + J_1$, se concluye que J es continua en todo punto.◄

► **Ejemplo 6.53** Sean $G = \{f \in C^n[a,b] \mid f(x) > 0\, \forall\, x \in [a,b]\}$ y $f_0 \in G$. Sea $\mu = \min_{a\leq x\leq b}\{f_0(x)\}$. Si $\|f - f_0\|_n < \mu$, entonces

$$|f(x) - f_0(x)| < \mu \quad \forall\, x \in [a,b].$$

Por tanto,

$$0 < f_0(x) - \mu < f(x) \quad \forall\, x \in [a,b];$$

lo cual implica $f \in G$; es decir, $B(f_0,\mu) \subset G$. Sea $J : G \to C[a,b]$ definido por $J(f) = 1/f$. Puesto que $\lim_{f\to f_0} f f_0 = f_0^2$ (cfr. el inciso $1d$ del teorema 6.7), existe $\delta > 0$ tal que

$$\|f f_0 - f_0^2\|_\infty < \frac{1}{2}m$$

si $0 < \|f - f_0\|_n < \delta$; donde $m = \text{mín}_{a \le x \le b}\{f_0^2(x)\} > 0$. Entonces, si $0 < \|f - f_0\|_n < \delta$,

$$-\frac{1}{2}m < f(x)f_0(x) - f_0^2(x) < \frac{1}{2}m$$

y, por tanto,

$$\begin{aligned} f(x)f_0(x) &> f_0^2(x) - \frac{1}{2}m \\ &\ge m - \frac{1}{2}m \\ &= \frac{1}{2}m \end{aligned}$$

para todo $x \in [a,b]$. Luego, si $0 < \|f - f_0\|_n < \delta$, para todo $x \in [a,b]$ se tiene

$$\begin{aligned} \left|\frac{1}{f(x)} - \frac{1}{f_0(x)}\right| &= \left|\frac{f_0(x) - f(x)}{f(x)f_0(x)}\right| \\ &\le \frac{2}{m}|f_0(x) - f(x)| \\ &\le \frac{2}{m}\|f_0 - f\|_n \end{aligned}$$

Por lo que

$$\left\|\frac{1}{f} - \frac{1}{f_0}\right\|_\infty \le \frac{2}{m}\|f_0 - f\|_n$$

y de donde se desprende

$$\lim_{f \to f_0} \frac{1}{f} = \frac{1}{f_0}.$$

Es decir, J es continuo en todo punto de G.◀

► **Ejemplo 6.54** **(Continuidad de las proyecciones)** Sea k un entero, con $1 \le k \le n$, se define la función $\pi_k : \mathbb{R}^n \to \mathbb{R}$,

$$\pi_k(\vec{x}) = x_k$$

para cada $\vec{x} = (x_1, x_2, \ldots, x_n)$. A esta función se le llama la k-ésima proyección del vector $\vec{x}$. Entonces, si $\vec{a} = (a_1, a_2, \ldots, a_n)$ es cualquier vector,

$$|\pi_k(\vec{x}) - \pi_k(\vec{a})| = |x_k - a_k| \le \|\vec{x} - \vec{a}\|_\infty$$

y, por tanto,

$$\lim_{\vec{x} \to \vec{a}} \pi_k(\vec{x}) = \pi_k(\vec{a}).$$

Esto implica que la función proyección π_k es continua en todo punto de $\mathbb{R}^n$, respecto a la norma cúbica y al valor absoluto; pero como en $\mathbb{R}^n$ todas las normas son equivalentes,[17] se concluye que π_k es continua en todo punto respecto a cualquier norma en $\mathbb{R}^n$.◀

[17] Cfr. teorema 4.22, pág. 327.

El hecho de que cualquier par de normas sean equivalentes en un espacio de dimensión finita (cfr. teorema 4.22), tiene una enorme ventaja; pues basta probar la continuidad respecto a una norma y entonces ésta también se cumple para cualquier otra norma. No haremos más estos comentarios cuando trabajemos en espacios de dimensión finita, simplemente utilizaremos la norma que más convenga y, entonces, los aspectos que se cumplan para ésta en cuanto a límites, continuidad y diferenciabilidad serán validos para todas las demás normas del espacio.

► **Ejemplo 6.55** **(Continuidad de los monomios)** Un monomio de varias variables es una expresión de la forma

$$Ax_1^{m_1}x_2^{m_2}\cdots x_n^{m_n}$$

donde A es una constante real y los números m_i son enteros no negativos. Si $\vec{x}=(x_1,x_2,\ldots,x_n)$ y se define

$$f(\vec{x})=Ax_1^{m_1}x_2^{m_2}\cdots x_n^{m_n}$$

entonces,

$$f(\vec{x})=A\pi_1(\vec{x})\pi_2(\vec{x})\cdots\pi_n(\vec{x}).$$

Dado que las proyecciones son continuas en todo punto y el producto de funciones continuas es también una función continua, se concluye que f es continua en todo punto de $\mathbb{R}^n$.◄

► **Ejemplo 6.56** **(Continuidad de polinomios de varias variables)** Un polinomio de varias variables es una suma finita de monomios de varias variables. Dado que la suma de funciones continuas da como resultado funciones continuas y los monomios de varias variables son funciones continuas en todo punto, se tiene que todo polinomio de varias variables, $p(\vec{x})$, es una función continua en todo punto.◄

► **Ejemplo 6.57** Sea $f:\mathbb{R}^3\to\mathbb{R}$ la función definida por

$$f(x,y,z)=\cos(x^2+2y-z)e^{xy}-\sqrt{x^2+y^2+z^4}$$

Entonces,

$$\cos(x^2+2y-z)=f_1(p_1(x,y,z))$$

donde $p_1(x,y,z)=x^2+2y-z$ y $f_1(t)=\cos(t)$. La primera función, p_1, es continua porque es un polinomio de varias variables, la segunda función es la función coseno que (como la mayoría sabe) es continua en todo punto. Puesto que la composición de funciones continuas en también continua, el primer factor del primer término define una función continua. De manera análoga, la función $(x,y,z)\mapsto e^{xy}$ es la composición de las funciones $(x,y,z)\mapsto xy$, monomio de varias variables, con la función continua de variable real $t\mapsto e^t$; por tanto, la función $(x,y,z)\mapsto\cos(x^2+2y-z)e^{xy}$ es continua en todo punto. La función $(x,y,z)\mapsto x^2+y^2+z^2\geq 0$ es continua en todo punto porque es un polinomio de varias variables; la función $t\mapsto\sqrt{t}$ es continua para todo $t\geq 0$; por tanto, la función $(x,y,z)\mapsto\sqrt{x^2+y^2+z^4}$ es continua en todo punto por ser el resultado de la composición de estas funciones. En consecuencia, al ser f suma de funciones continuas en todo punto, f es una función continua en todo punto.◄

El propósito de lo precedente es mostrar al lector cómo es posible distinguir rápidamente por inspección cuándo una función de varias variables con valores reales es continua. Éste en realidad es un

proceso sencillo pero muy importante para aplicar, como paso intermedio, resultados que veremos más adelante. El siguiente teorema es fácil de probar y su demostración se deja de ejercicio al lector.

Teorema 6.9 *Sean* $\mathbf{E}$ *un espacio vectorial y* $J : D \subset \mathbf{E} \to \mathbb{R}^n$ *una función; con funciones componentes (o coordenadas)* $f_i : D \to \mathbb{R}^n$*; es decir,*

$$J(\vec{u}) = (f_1(\vec{u}), f_2(\vec{u}), \ldots, f_n(\vec{u})).$$

1. *Si* $\vec{l} = (l_1, l_2, \ldots, l_n)$*, entonces*

$$\lim_{\vec{u}\to\vec{u}_0} J(\vec{u}) = \vec{l} \Leftrightarrow \lim_{\vec{u}\to u_0} f_i(\vec{u}) = l_i \quad \forall\, i.$$

2. J *es continua en* $\vec{u}_0$*, si y sólo si* f_i *es continua en* $\vec{u}_0$ *para todo i.*

▶ **Ejemplo 6.58** Sea $J : \mathbb{R}^3 \to \mathbb{R}^2$, definida por $J(x,y,z) = (xy^3 + z^2, xe^y)$. En este caso $f_1(x,y,z) = xy^3 + z^2$ y $f_2(x,y,z) = xe^y$ son las funciones componentes de J. Por argumentos similares a los expuestos en los ejemplos previos (el lector debería intentar escribirlos) se comprueba que f_1 y f_2 son funciones continuas en todo punto. Por tanto, J es continua en todo punto de $\mathbb{R}^3$.◀

La continuidad de una función es una propiedad puntual; es decir, una función puede ser continua en un punto y en otros no. Sin embargo, cuando la función es lineal la situación cambia; si es continua en un punto, entonces va a ser continua en todos los demás y, por tanto, si es discontinua en un punto, también es discontinua en todos los demás. Hacemos patente esta característica fundamental de las funciones lineales en el siguiente teorema; que además da lugar a definir una norma en el espacio $\mathscr{L}(\mathbf{E},\mathbf{F})$.

Teorema 6.10 *Sean* $(\mathbf{E}, \|\cdot\|)$, $(\mathbf{F}, \|\cdot\|)$ *espacios vectoriales normados y* $J : \mathbf{E} \to \mathbf{F}$ *una transformación lineal, las tres condiciones siguientes son equivalentes a pares:*

1. J *es continua en* $0 \in \mathbf{E}$
2. *Existe* $M > 0$ *tal que*

$$\|J(f)\| \leq M\,\|f\| \;\; \forall f \in \mathbf{E}$$

3. J *es continua*[18] *en todo punto de* $\mathbf{E}$.

DEMOSTRACIÓN ■ 1)⇒ 2) Sea $\delta > 0$ tal que $\|g\| \leq \delta \Rightarrow \|J(g)\| \leq 1$. Sean $f \in \mathbf{E} - \{0\}$ y $g = \dfrac{\delta}{\|f\|} f$, entonces $\|g\| \leq \delta$ y, por tanto,

$$\begin{aligned} 1 &\geq \|J(g)\| \\ &= \left\| J\left(\frac{\delta}{\|f\|} f\right) \right\| \\ &= \frac{\delta}{\|f\|} \|J(f)\| \end{aligned}$$

[18] Debido a la equivalencia de las condiciones 2 y 3 de este teorema, se acostumbra decir que una función lineal continua entre espacios vectoriales es una aplicación lineal acotada.

de donde

$$\|J(f)\| \leq M\|f\| \ \forall f \in \mathbf{E}$$

con $M = \dfrac{1}{\delta}$.

2) $\Rightarrow$ 3) Sean $f_0, f \in \mathbf{E}$, entonces

$$\begin{aligned} \|J(f) - J(f_0)\| &= \|J(f - f_0)\| \\ &\leq M\|f - f_0\| \end{aligned}$$

de donde se tiene que la siguiente implicación es válida

$$\|f - f_0\| < \delta \Rightarrow \|J(f) - J(f_0)\| < \varepsilon$$

tomando $\delta = \varepsilon/(M+1)$.

3) $\Rightarrow$ 1) Es obvia. ■

En síntesis, el teorema anterior indica que una aplicación lineal cuando es continua en un punto es continua en todo el espacio. Por esta razón afirmaremos simplemente que una aplicación lineal es continua, quedando implícitamente entendido que es continua en todo su espacio de definición. Por otra parte, es muy conveniente tener presente la condición número 2 del teorema precedente como una condición necesaria y suficiente para que una aplicación lineal sea continua. Como mencionamos antes, se puede definir una norma en el espacio de transformaciones lineales continuas. Antes de continuar establecemos unos cuantos preliminares técnicos.

1. Si $A \subset \mathbb{R}$ y existe M tal que $M \leq a \ \forall \ a \in A$, entonces se dice que M es una cota inferior de A y que A está acotado inferiormente.
2. Un principio fundamental de los números reales, es el que establece que todo conjunto inferiormente acotado posee una máxima cota inferior; es decir, existe $M_0 \in \mathbb{R}$, cota inferior de A, tal que $M \leq M_0$ para toda cota inferior M de A. A M_0 se le llama el ínfimo de A y se representa por el símbolo $\text{ínf}(A)$.

Teorema 6.11 *Sean* $\mathbf{E}$ *y* $\mathbf{F}$ *un par de espacios vectoriales normados, entonces:*

1. *El conjunto*

$$\mathscr{L}(\mathbf{E}, \mathbf{F}) = \{J : \mathbf{E} \to \mathbf{F} \,|\, J \textit{ es lineal y continua}\}$$

es un espacio con la suma usual de funciones y el producto estándar de un escalar por una función.

2. *La aplicación* $\|\cdot\| : \mathscr{L}(\mathbf{E}, \mathbf{F}) \to \mathbb{R}$ *definida por*

$$\|J\| = \text{ínf}\{M > 0 \,|\, \|J(f)\| \leq M\|f\| \quad \forall f \in \mathbf{E}\}$$

es una norma.

3. $$\|J(f)\| \leq \|J\|\,\|f\| \ \forall f \in \mathbf{E}.$$

Cuando el espacio $\mathbf{E}$ sea de dimensión finita, entonces cualquier transformación lineal $J:\mathbf{E}\to\mathbf{F}$ es continua. Efectivamente, sean $\{\varphi_1,\varphi_2\ldots,\varphi_n\}$ una base del espacio $\mathbf{E}$, y $f=\alpha_1\varphi_1+\alpha_2\varphi_2+\cdots+\alpha_n\varphi_n$ cualquier vector del espacio $\mathbf{E}$, entonces

$$\begin{aligned}\|J(f)\| &= \left\|\sum_{i=1}^{n} J(\alpha_i\varphi_i)\right\| \\ &= \left\|\sum_{i=1}^{n} \alpha_i J(\varphi_i)\right\| \\ &\le \sum_{i=1}^{n} |\alpha_i|\,\|J(\varphi_i)\| \\ &\le \sum_{i=1}^{n} \|f\|_\infty\,\|J(\varphi_i)\| \\ &= \|f\|_\infty \sum_{i=1}^{n} \|J(\varphi_i)\| \\ &= M\,\|f\|_\infty\end{aligned}$$

con $M=\sum_{i=1}^{n}\|J(\varphi_i)\|$. Se deduce, de la condición 2 del teorema 6.10, que J es continua; pues todas las normas en espacios de dimensión finita son equivalentes. Sin embargo, se puede probar que si $\mathbf{E}$ tiene dimensión infinita y $\mathbf{F}$ es cualquier espacio, siempre existe una transformación lineal discontinua $L:\mathbf{E}\to\mathbf{F}$.

Diferenciabilidad

Recordemos que una función J de una variable real con valores en $\mathbb{R}$ es derivable en el punto x_0 si el límite

$$\lim_{h\to 0}\frac{J(x_0+h)-J(x_0)}{h}=J'(x_0) \tag{6.62}$$

existe. A $J'(x_0)$ se le llama la derivada de la función J en el punto x_0 y, geométricamente, es la pendiente de la recta tangente a la gráfica de la función J en el punto x_0. Notemos que (6.62) es equivalente a

$$\lim_{h\to 0}\frac{J(x_0+h)-J(x_0)-J'(x_0)h}{h}=0$$

Entonces, para $\varepsilon=1$ existe $\delta>0$ tal que

$$|h|<\delta \Rightarrow |J(x_0+h)-(J'(x_0)h+J(x_0))|<|h|$$

Esto significa que si en lugar de evaluar la función J en x_0+h se evalúa en la recta tangente $T(x)=J'(x_0)(x-x_0)+J(x_0)$ se obtiene

$$|J(x_0+h)-T(x_0+h)|=|J(x_0+h)-(J'(x_0)h+J(x_0))|<h$$

Es decir, el error cometido al hacer esta aproximación es aún menor que el propio h. Así, si h es pequeño, digamos menor a 10^{-6}, entonces $|J(x_0+h)-T(x_0+h)|$ es todavía más pequeño. Por tanto, la derivada da la posibilidad de que mediante la recta tangente se obtenga una aproximación lineal de un orden

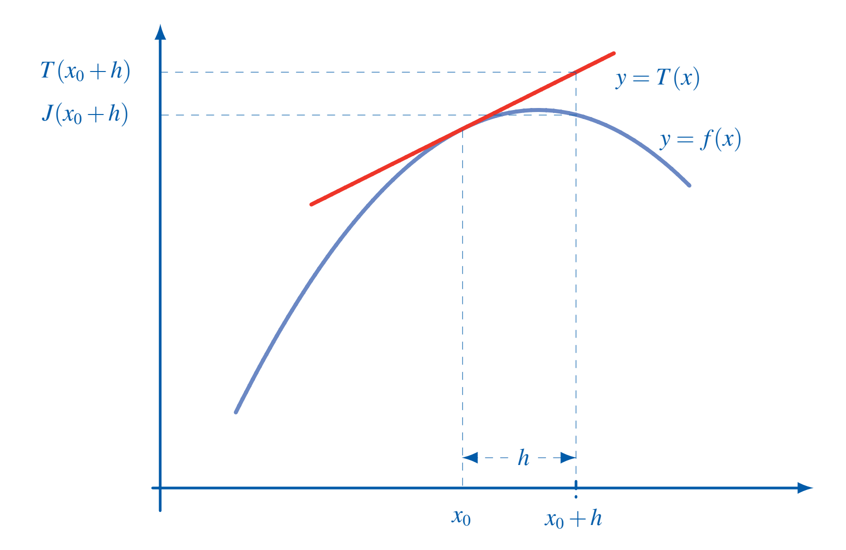


Figura 6-30 •

menor que el incremento de la variable independiente, para valores cercanos al punto donde se calcula la derivada; es decir

$$J(x_0+h)\approx J'(x_0)h+J(x_0)=T(x_0+h)$$

si h es pequeño. Por tanto, la recta tangente es la mejor estimación lineal que tiene la función cerca del punto de tangencia (cfr. figura 6-30). Por otra parte, note que la aplicación $L(h)=J'(x_0)h$ es una transformación lineal del espacio $\mathbb{R}$ en el espacio $\mathbb{R}$ y $T(h)=L(h)+J(x_0)$. Motivados por las ideas precedentes damos a continuación la generalización de la derivación, como el proceso que da por resultado la mejor estimación lineal a una función cerca del punto donde se obtiene la derivada.

❍ Nota 6.5

1. Antes de continuar recordemos que si L es una transformación lineal, se escribe con frecuencia Lu en lugar de $L(u)$. Utilizaremos las dos notaciones según convenga.
2. Siempre que tratemos el tema de diferenciabilidad, concepto que daremos en seguida, de una función en un punto f_0 supondremos, aunque no se mencione explícitamente, que existe un $r>0$ tal que la función en cuestión está definida en todos los puntos de la bola abierta $B(f_0,r)$.

Definición 6.10 *Sean* $(\mathbf{E},\|\cdot\|)$ *y* $(\mathbf{F},\|\cdot\|)$ *dos espacios vectoriales normados y* $J:\mathbf{E}\to\mathbf{F}$ *una función. Se dice que* J *es diferenciable en* f_0 *si existe una transformación lineal continua,* $L(f_0):\mathbf{E}\to\mathbf{F}$*, tal que*[19]

$$\lim_{h\to 0}\frac{J(f_0+h)-J(f_0)-L(f_0)h}{\|h\|}=0 \qquad (6.63)$$

Se puede probar que de ser diferenciable J *en* f_0 *la aplicación lineal* $L(f_0)$ *que satisface (6.63) es única. Se le llama la diferencial de* J *en* f_0 *y se denota por* $dJ(f_0)$ *o* $J'(f_0)$.

[19] Aquí $h\to 0$ significa que el elemento h tiende al neutro aditivo 0 del espacio vectorial $\mathbf{E}$; mientras que el resultado del límite representa el neutro aditivo del espacio $\mathbf{F}$.

Es pertinente tener muy claros los siguientes puntos:

1. Los valores de la función J son vectores en el espacio $\mathbf{F}$, por tanto, el resultado de la operación indicada en el numerador de (6.63) es un vector de $\mathbf{F}$; mientras que $\|h\|$ es un escalar. Por eso, en realidad, la notación
$$\frac{J(f_0+h)-J(x_0)-L(f_0)h}{\|h\|}$$
significa
$$\frac{1}{\|h\|}\left(J(f_0+h)-J(x_0)-L(f_0)h\right).$$
2. La diferencial es en sí una función, una transformación del espacio $\mathbf{E}$ en el espacio $\mathbf{F}$; no es un vector o un escalar. Por eso la notación provisional $L(f_0)$ y las notaciones que usaremos para representar la diferencial, $J'(f_0)$ y $dJ(f_0)$, representan cada una de ellas una aplicación lineal y no la evaluación de una función en un argumento. Sin embargo, la diferencial es otra función cuando se cambia f_0; es decir, $J'(f_0)$ es una función y $J'(f_1)$, en general, es otra transformación cuando $f_1 \neq f_0$. Por eso es que la notación $J'(f_0)$ hace énfasis que la diferencial de una función depende del valor f_0.
3. La notación $L(f_0)h$ es la evaluación de la diferencial en h; esto es $L(f_0)h = L(f_0)(h)$, cuyo resultado sí es un vector, un elemento de $\mathbf{F}$.
4. Por el mismo argumento dado antes, para el caso de funciones reales de una variable real (6.63), implica
$$\|(J(f_0+h)-J(f_0))-J'(f_0)h\| < \|h\|$$
para $\|h\|$ suficientemente pequeña. Es decir,
$$J(f_0+h)-J(x_0) \approx J'(f_0)h$$
si $\|h\|$ es pequeña. Esto es, la diferencia $J(f_0+h)-J(f_0)$ se puede aproximar con $J'(f_0)h$; ésta es la razón del sustantivo "diferencial" empleado para la transformación lineal $J'(f_0)$.
5. Dado que el límite de un operador es invariante bajo el cambio de normas equivalentes, la condición de diferenciabilidad se preserva si se cambian las normas originales en los espacios $\mathbf{E}$ y $\mathbf{F}$ por normas equivalentes y la diferencial del operador es la misma en ambos casos.

► **Ejemplo 6.59** En el caso particular de que $\mathbf{E} = \mathbb{R}$ y $J : \mathbb{R} \to \mathbb{R}$, sea una función derivable en x_0, entonces la función $h \mapsto J'(x_0)\,h$ es lineal y continua; además, por la definición de derivada de funciones reales de una variable, se tiene
$$\lim_{h\to 0}\frac{J(x_0+h)-J(x_0)-J'(x_0)\,h}{h} = 0$$
Por lo que J es diferenciable en el sentido de la precedente definición y $dJ(x_0)\,h = J'(x_0)\,h$. Es decir, la definición 6.10 tiene como caso particular la definición de derivada de cursos elementales de cálculo diferencial.◄

▶ **Ejemplo 6.60** **(Diferenciabilidad de las transformaciones lineales)** Sean $(\mathbf{E}, \|\cdot\|)$ y $(\mathbf{F}, \|\cdot\|)$ un par de espacios vectoriales normados y $L : \mathbf{E} \to \mathbf{F}$ una transformación lineal continua, entonces L es diferenciable en todo punto $f \in \mathbf{E}$ y además $dL(f) = L \; \forall f \in \mathbf{E}$. Efectivamente:

$$\begin{aligned} \lim_{h\to 0} \frac{L(f+h)-L(f)-Lh}{\|h\|} &= \lim_{h\to 0} \frac{L(f)+Lh-L(f)-Lh}{\|h\|}. \\ &= 0 \blacktriangleleft \end{aligned}$$

▶ **Ejemplo 6.61** **(Diferenciabilidad del operador integración)** Sea $J : C[a,b] \to \mathbb{R}$ el funcional definido por $J(f) = \int_a^b f$. Claramente J es lineal y si $g \in C[a,b]$,

$$\begin{aligned} |J(f) - J(g)| &= \left| \int_a^b f - \int_a^b g \right| \\ &= \left| \int_a^b (f-g) \right| \\ &\le \int_a^b |f-g| \\ &\le \int_a^b \|f-g\|_\infty \\ &= (b-a)\|f-g\|_\infty \end{aligned}$$

de donde se deduce que J es continuo. Por el ejemplo anterior J es diferenciable y

$$dJ(f)h = J(h) = \int_a^b h \blacktriangleleft$$

Como en el caso de funciones de una variable real y valores reales la diferenciabilidad implica continuidad.

Teorema 6.12 *Sean $J : D \subset \mathbf{E} \to \mathbf{F}$ una función, $f_0 \in \mathbf{E}$.*

1. *Si J es diferenciable en f_0, entonces J es continua en f_0.*
2. *Si J es continua en f_0 y existe una transformación lineal $L : \mathbf{E} \to \mathbf{F}$ tal que*

$$\lim_{h\to 0} \frac{J(f_0+h) - J(f_0) - Lh}{\|h\|} = 0 \tag{6.64}$$

entonces la transformación lineal L es continua. Por tanto, J es diferenciable en f_0 y $L = J'(f_0)$.

DEMOSTRACIÓN ■ 1. Sea $T : \mathbf{E} - \{0\} \to \mathbf{F}$, definida por

$$T(h) = \begin{cases} \dfrac{1}{\|h\|}\left(J(f_0+h) - J(f_0) - J'(f_0)h\right), & \text{si } h \neq 0 \\ 0, & \text{si } h = 0 \end{cases}.$$

Entonces,

$$\lim_{h\to 0} T(h) = 0.$$

Ya que

$$J(f_0+h)-J(f_0)=J'(f_0)h+\|h\|\,T(h)$$

se tiene

$$\begin{aligned}\lim_{h\to 0}\left[J(f_0+h)-J(f_0)\right] &= \lim_{h\to 0}\left(J'(f_0)h+\|h\|\,T(h)\right)\\ &= 0+0(0)\\ &= 0\end{aligned}$$

y, por tanto, J es continua en f_0.

2. Supongamos ahora que J es continua en f_0 y que se satisface (6.64); sea

$$T(h)=\begin{cases}\frac{1}{\|h\|}\left(J(f_0+h)-J(f_0)-Lh\right), & \text{si } h\neq 0\\ 0 & \text{si } h=0\end{cases}$$

Entonces,

$$\lim_{h\to 0}T(h)=0$$

ya que

$$Lh=J(f_0+h)-J(f_0)-\|h\|\,T(h)$$

se tiene

$$\begin{aligned}\lim_{h\to 0}Lh &= \lim\left[J(f_0+h)-J(f_0)-\|h\|\,T(h)\right]\\ &= 0-0\\ &= 0.\end{aligned}$$

Por tanto, J es diferenciable y, por la unicidad de la diferencial, $L=J'(f_0)$. ■

El teorema anterior es muy útil, pues con frecuencia por simple inspección se puede verificar la continuidad de la función J y, entonces, resta únicamente probar que se cumpla la condición (6.64) para que J sea diferenciable en f_0 y que, por tanto, $J'(f_0)=L$.

El siguiente paso consiste en determinar la diferencial de una función. El teorema 6.13 indica de qué forma debe ser la diferencial de una función en el caso de que ésta sea diferenciable.

Teorema 6.13 *Sean* $(\mathbf{E},\|\cdot\|)$ *y* $(\mathbf{F},\|\cdot\|)$ *espacios vectoriales normados y* $J:\mathbf{E}\to\mathbf{F}$ *una función diferenciable en* $f_0\in\mathbf{E}$. *Sea* $u\in\mathbf{E}$, *entonces*

$$\lim_{\alpha\to 0}\frac{J(f_0+\alpha u)-J(f_0)}{\alpha}=dJ(f_0)\,u \qquad (6.65)$$

DEMOSTRACIÓN ■ Supongamos primero que $\|u\|=1$. Sea $h=\alpha u$, entonces $h\to 0$ cuando $\alpha\to 0$; luego

$$
\begin{aligned}
\lim_{\alpha\to 0}\frac{J(f_0+\alpha u)-J(f_0)-\alpha dJ(f_0)u}{\alpha} &= \lim_{\alpha\to 0}\frac{J(f_0+\alpha u)-J(f_0)-dJ(f_0)(\alpha u)}{\alpha}\\
&= \pm\lim_{h\to 0}\frac{J(f_0+h)-J(f_0)-dJ(f_0)(h)}{\|h\|}\\
&= 0
\end{aligned}
$$

de donde se cumple (6.65). Ahora supongamos que u es un elemento cualquiera de $\mathbf{E}-\{0\}$. Sea $v=(1/\|u\|)\,u$, entonces v es unitario, así que

$$
\begin{aligned}
\frac{1}{\|u\|}dJ(f_0)u &= dJ(f_0)v\\
&= \lim_{t\to 0}\frac{J(f_0+tv)-J(f_0)}{t}\\
&= \lim_{t\to 0}\frac{J\left(f_0+\frac{t}{\|u\|}u\right)-J(f_0)}{t}\\
&= \frac{1}{\|u\|}\lim_{t\to 0}\frac{J\left(f_0+\frac{t}{\|u\|}u\right)-J(f_0)}{\frac{t}{\|u\|}}\\
&= \frac{1}{\|u\|}\lim_{\alpha\to 0}\frac{J(f_0+\alpha u)-J(f_0)}{\alpha} \qquad (\text{donde } \alpha=\frac{t}{\|u\|}).
\end{aligned}
$$

De donde nuevamente se tiene (6.65). ■

Definición 6.11 **(Derivada direccional)** *Al límite en el lado izquierdo de (6.65) se le llama la derivada direccional de J en f_0 en la dirección del vector u y se representa como*

$$\partial_u J(f_0)$$

siempre que dicho límite exista.

► **Ejemplo 6.62** Sea $J: C^1[a,b]\to C[a,b]$ la aplicacion definida como $J(f)=1+[f']^2$. Entonces,

$$
\begin{aligned}
\lim_{t\to 0}\frac{J(f+tu)-J(f)}{t} &= \lim_{t\to 0}\frac{1+[f']^2+2tf'u'+t^2[u']^2-1-[f']^2}{t}\\
&= 2f'u'.
\end{aligned}
$$

Sea $dJ(f): C^1[a,b]\to C[a,b]$ definida como $dJ(f)u=2f'u'$. Entonces claramente $dJ(f)$ es lineal y

$$
\begin{aligned}
\|dJ(f)u\|_\infty &= \|2f'u'\|_\infty\\
&\le 2\|f'\|_\infty\|u'\|_\infty\\
&\le 2\|f'\|_\infty\|u\|_1
\end{aligned}
$$

de donde se infiere que $dJ(f)$ es continua. Además,

$$\begin{aligned}\lim_{h\to 0}\frac{J(f+h)-J(f)-dJ(f)h}{\|h\|_1} &= \lim_{h\to 0}\frac{1+[f']^2+2f'h'+[h']^2-1-[f']^2-2f'h'}{\|h\|_1}\\ &= \lim_{h\to 0}\frac{[h']^2}{\|h\|_1}\\ &= 0\end{aligned}$$

pues

$$0\le \lim_{h\to 0}\frac{\left\|[h']^2\right\|_\infty}{\|h\|_1}\le \lim_{h\to 0}\frac{\|h\|_1^2}{\|h\|_1}=\lim_{h\to 0}\|h\|_1=0$$

Lo cual prueba que J es diferenciable y

$$dJ(f)h=2f'h'.◀$$

▶ **Ejemplo 6.63** Sea $G=\left\{f\in C^1[a,b]\,|\,f(x)>0\;\forall\,x\in[a,b]\right\}$. Por el ejemplo 6.53, dado cualquier $f\in G$, existe $r>0$ tal que $B(f,r)\subset G$. Sea $J:G\to C[a,b]$ el operador definido por $J(f)=\sqrt{f}$, entonces

$$\begin{aligned}\lim_{t\to 0}\frac{J(f+tu)-J(f)}{t} &= \lim_{t\to 0}\frac{\sqrt{f+tu}-\sqrt{f}}{t}\\ &= \lim_{t\to 0}\frac{\sqrt{f+tu}-\sqrt{f}}{t}\cdot\frac{\sqrt{f+tu}+\sqrt{f}}{\sqrt{f+tu}+\sqrt{f}}\\ &= \lim_{t\to 0}\frac{f+tu-f}{t\left(\sqrt{f+tu}+\sqrt{f}\right)}\\ &= \lim_{t\to 0}\frac{u}{\sqrt{f+tu}+\sqrt{f}}=\frac{u}{2\sqrt{f}}\;;\end{aligned}$$

pues por el ejemplo 6.50 y el inciso 1*(a)* del teorema 6.7, $\lim_{t\to 0}\sqrt{f+th}=\sqrt{f}$ y, por tanto, $\lim_{t\to 0}\sqrt{f+th}+\sqrt{f}=2\sqrt{f}$ (cfr. ejemplo 6.49 e inciso 1*(a)* del teorema 6.7); entonces, del ejemplo 6.53 y el inciso 2 del teorema 6.7, se tiene $\lim_{t\to 0}1/(\sqrt{f+th}+\sqrt{f})=1/2\sqrt{f}$; y por la propiedad 6.7 del teorema 6.7 y el ejemplo 6.49, $\lim_{t\to 0}u/(\sqrt{f+th}+\sqrt{f})=u/2\sqrt{f}$.

Sea $dJ(f):C^1[a,b]\to C[a,b]$ definido por $dJ(f)u=\dfrac{u}{2\sqrt{f}}$. Claramente $dJ(f)$ es lineal y $\|dJ(f)u\|_\infty=\left\|\dfrac{u}{2\sqrt{f}}\right\|_\infty$. Sea $k=\min_{a\le x\le b}2\sqrt{f(x)}$, entonces $\left\|\dfrac{u}{2\sqrt{f}}\right\|_\infty\le\dfrac{1}{k}\|u\|_1$; es decir, $dJ(f)$ es continuo. Finalmente,

$$\begin{aligned}\lim_{h\to 0}\frac{J(f+h)-J(f)-dJ(f)h}{\|h\|_1} &= \lim_{h\to 0}\frac{\sqrt{f+h}-\sqrt{f}-\dfrac{h}{2\sqrt{f}}}{\|h\|_1}\\ &= \lim_{h\to 0}\frac{\left(\sqrt{f+h}-\sqrt{f}-\dfrac{h}{2\sqrt{f}}\right)\left(\sqrt{f+h}+\sqrt{f}\right)}{\|h\|_1\left(\sqrt{f+h}+\sqrt{f}\right)}\\ &= \lim_{h\to 0}\frac{f+h-f-\dfrac{h}{2\sqrt{f}}\left(\sqrt{f+h}+\sqrt{f}\right)}{\|h\|_1\left(\sqrt{f+h}+\sqrt{f}\right)}\\ &= \lim_{h\to 0}\frac{2\sqrt{f}h-h\left(\sqrt{f+h}+\sqrt{f}\right)}{2\sqrt{f}\,\|h\|_1\left(\sqrt{f+h}+\sqrt{f}\right)}\\ &= -\frac{1}{2\sqrt{f}}\lim_{h\to 0}\frac{h\left(\sqrt{f+h}-\sqrt{f}\right)}{\|h\|_1\left(\sqrt{f+h}+\sqrt{f}\right)}=0\end{aligned}$$

pues

$$\lim_{h\to 0}\left\|\frac{h\left(\sqrt{f+h}-\sqrt{f}\right)}{\|h\|_1\left(\sqrt{f+h}+\sqrt{f}\right)}\right\|_\infty \le \lim_{h\to 0}\frac{\|h\|_\infty}{\|h\|_1}\left\|\frac{\sqrt{f+h}-\sqrt{f}}{\sqrt{f+h}+\sqrt{f}}\right\|_\infty$$
$$= \lim_{h\to 0}\left\|\frac{\sqrt{f+h}-\sqrt{f}}{\sqrt{f+h}+\sqrt{f}}\right\|_\infty$$
$$= 0$$

ya que si $m = \mathrm{mín}_{a\le x\le b}\sqrt{f(x)}$, entonces

$$\left|\sqrt{f(x)+h(x)}-\sqrt{f(x)}\right| = \frac{|h(x)|}{\sqrt{f(x)+h(x)}+\sqrt{f(x)}}$$
$$\le \frac{|h(x)|}{\left|\sqrt{f(x)}\right|}$$
$$\le \frac{\|h\|_1}{m}$$

por lo que

$$\left\|\frac{\sqrt{f+h}-\sqrt{f}}{\sqrt{f+h}-\sqrt{f}}\right\| \le \frac{\|h\|_1}{m^2}\,.$$

Con lo cual queda demostrado que

$$dJ(f) = \frac{u}{2\sqrt{f}} \quad \forall u \in G.◀$$

▶ **Ejemplo 6.64** Sean G el conjunto definido en el ejemplo anterior y $J : G \subset C^1[a,b] \to C[a,b]$ el operador definido por $J(f) = 1/f$. Mostrar que J es diferenciable en todo punto $f \in G$ y encontrar $J'(f)$.◀

Solución Si $h \in C[a,b]$, entonces

$$\lim_{t\to 0}\frac{J(f+th)-J(f)}{t} = \lim_{t\to 0}\frac{\frac{1}{f+th}-\frac{1}{f}}{t}$$
$$= \lim_{t\to 0}\frac{\frac{-th}{f(f+th)}}{t}$$
$$= \lim_{t\to 0}\frac{-h}{f(f+th)}$$
$$= -\frac{h}{f^2}$$

(el lector debe justificar por su cuenta, como hicimos en los ejemplos anteriores, la validez de la última igualdad). Sea $L : C^1[a,b] \to C[a,b]$ definido por $L(h) = -h/f^2$; claramente L es lineal. Dado que $f \in C^1[a,b]$, existe $m_0 = \mathrm{mín}_{a\le x\le b}\{f^2(x)\} > 0$, y, por tanto,

$$\left|1/f^2(x)\right| \le m_0$$

para todo $x \in [a,b]$; luego

$$\left|-\frac{h(x)}{f^2(x)}\right| \le m_0\,|h(x)| \le m_0\,\|h\|$$

para todo $x \in [a,b]$. Entonces

$$\left\| \frac{h}{f^2} \right\| \leq m_0 \|h\|$$

para todo $h \in C[a,b]$. Lo cual prueba que L es continua. Finalmente,

$$\begin{aligned}
\lim_{h\to 0} \frac{J(f+h)-J(f)-Lh}{\|h\|} &= \lim_{h\to 0} \frac{\frac{1}{f+h}-\frac{1}{f}+\frac{h}{f^2}}{h} \\
&= \lim_{h\to 0} \frac{\frac{f^2-f(f+h)+h(f+h)}{f^2(f+h)}}{h} \\
&= \lim_{h\to 0} \frac{h^2}{h(f+h)} \\
&= \lim_{h\to 0} \frac{h}{f+h} \\
&= \frac{0}{f} \\
&= 0.
\end{aligned}$$

(nuevamente queda de ejercicio al lector justificar rigurosamente la penúltima igualdad). Por tanto, J es diferenciable en todo punto de G y $J'(f)h = -h/f^2$. ✔

Hemos visto, en los ejemplos previos y lo que les precede, que una manera de conocer la diferencial de una función, es por medio de la derivada direccional, y que toda función diferenciable tiene derivada direccional en la dirección de cualquier vector; sin embargo, se pueden dar ejemplos de funciones que en un punto dado tienen derivada direccional en cualquier dirección, pero no son diferenciables en éste. El siguiente teorema establece condiciones suficientes para que una función sea diferenciable utilizando información de las derivadas direccionales.

Teorema 6.14 **(Condiciones suficientes para diferenciabilidad)** *Sean* $\mathbf{E}$, $\mathbf{F}$ *dos espacios normados y* $J : \mathbf{E} \to \mathbf{F}$ *un operador. Supongamos que en* f_0 *se cumplen las siguientes condiciones:*

1. *La derivada en la dirección de cualquier vector* $h \in \mathbf{E}$ *existe en todos los puntos de un conjunto abierto* G *que contiene a* f_0.
2. *La aplicación* $h \mapsto \partial_h(u)$ *es lineal y continua para cada* $u \in G$.
3. *La aplicación* $u \mapsto \Phi(u) \in \mathscr{L}(\mathbf{E},\mathbf{F})$, *donde* $\Phi(u)h = \partial_h(u)$, *es continua en* f_0.

Entonces, J *es diferenciable en* f_0 *y además,* $J'(f_0)h = \partial_h(f_0)$.

Como en el caso de una variable real, se tiene una regla para calcular la diferencial de la composición de funciones. Enunciamos, sin probar, el siguiente teorema e invitamos al lector a que intente su demostración.

Teorema 6.15 **(Regla de la cadena)** *Sean* $(\mathbf{E}, \|\cdot\|)$, $(\mathbf{F}, \|\cdot\|)$ *y* $(\mathbf{G}, \|\cdot\|)$ *espacios vectoriales;* $J_1 : \mathbf{E} \to \mathbf{F}$, $J_2 : \mathbf{F} \to \mathbf{G}$ *un par de funciones. Si* J_1 *es diferenciable en* $f \in \mathbf{E}$ *y* J_2 *es diferenciable en* $J_1(f) \in \mathbf{F}$, *entonces la aplicación* $J = J_2 \circ J_1 : \mathbf{E} \to \mathbf{G}$ *es diferenciable en* f *y, además,*

$$dJ(f) = dJ_2(J_1(f)) \circ dJ_1(f)$$

► **Ejemplo 6.65** Sea $J : C^1[a,b] \to C[a,b]$ la función definida como $J(f) = \sqrt{1+[f']^2}$. Entonces $J = J_2 \circ J_1$ donde $J_1(f) = 1+[f']^2$ y $J_2(g) = \sqrt{g}$ son las transformaciones de los ejemplos 6.63 y 6.62; por tanto, J es diferenciable y

$$\begin{aligned} dJ(f) &= dJ_2(J_1(f)) \circ dJ_1(f) \\ dJ(f)\vec{h} &= (dJ_2(J_1(f)) \circ dJ_1(f))h \\ &= dJ_2(J_1(f))(dJ_1(f)h) \\ &= dJ_2(J_1(f))(2f'h') \\ &= dJ_2\left(1+[f']^2\right)(2f'h') \\ &= \frac{f'h'}{\sqrt{1+[f']^2}}. \end{aligned}$$ ◄

► **Ejemplo 6.66** Sea $J : C^1[a,b] \to \mathbb{R}$ el funcional definido como

$$J(f) = \int_a^b \sqrt{1+[f']^2}$$

el operador que asigna a cada función derivable en $[a,b]$ la longitud de la curva $y = f(x)$ en este intervalo. Entonces, $J = J_2 \circ J_1$, donde $J_1 : C^1[a,b] \to C[a,b]$ es la aplicación $J_1(f) = \sqrt{1+[f']^2}$; y $J_2 : C[a,b] \to \mathbb{R}$ está definida por $J(g) = \int_a^b g$. Por los ejemplos 6.61 y 6.65 y la regla de la cadena:

$$\begin{aligned} dJ(f)h &= (dJ_2(J_1(f)) \circ dJ_1(f))h \\ &= dJ_2(J_1(f))(dJ_1(f)h) \\ &= \int_a^b dJ_1(f)h \\ &= \int_a^b \frac{f'h'}{\sqrt{1+[f']^2}}. \end{aligned}$$ ◄

La siguiente proposición es una generalización de la regla del producto del cálculo diferencial de una variable; su demostración se deja de ejercicio al lector.

Teorema 6.16 **(Regla del producto)** *Sean $J_1, J_2 : C[a,b] \to C[a,b]$ un par de funciones diferenciables en $f_0 \in C[a,b]$. Entonces, la función $J_1J_2 : C[a,b] \to C[a,b]$ definida por $(J_1J_2)(f) = J_1(f)J_2(f)$, para cada $f \in C[a,b]$, es diferenciable en f_0 y*

$$(J_1J_2)'(f_0)h = J_1'(f_0)hJ_2(f_0) + J_1(f_0)J_2'(f_0)h$$

$\forall h \in C[a,b]$.

► **Ejemplo 6.67** Sean G el conjunto del ejemplo 6.63, $J : G \subset C^1[a,b] \to C[a,b]$ la función definida por $J(f) = \sqrt{\frac{1+(f')^2}{f}}$. Entonces, $J(f) = J_1(f)J_2(f)$ donde $J_1(f) = \sqrt{1+(f')^2}$ y $J_2(f) = \frac{1}{\sqrt{f}}$. Por el ejemplo 6.65,

$$J_1'(f)h = \frac{f'h'}{\sqrt{1+(f')^2}};$$

mientras que por los ejemplos 6.63 y 6.64 y la regla de la cadena,

$$J_2'(f)h = \sqrt{1+(f')^2}\left(-\frac{h}{2f^{3/2}}\right).$$

Entonces, por la regla del producto

$$\begin{aligned} J'(f)h &= J_1'(f)hJ_2(f) + J_1(f)J_2'(f)h \\ &= \frac{f'h'}{f^{1/2}\left(1+(f')^2\right)^{1/2}} - \frac{\left(1+(f')^2\right)^{1/2}h}{2f^{3/2}}. \blacktriangleleft \end{aligned}$$

6.8.3 *Cálculo diferencial para funcionales en* $\mathbb{R}^n$

Un caso particularmente importante del cálculo diferencial en espacios vectoriales, es cuando se tiene una función f con dominio D contenido en $\mathbb{R}^n$ y valores en $\mathbb{R}$, un **funcional en** $\mathbb{R}^n$. A su estudio nos abocamos a contiuación .

Supongamos que $f : D \subset \mathbb{R}^n \to \mathbb{R}$ es un funcional, $y = f(x_1, x_2, \ldots, x_n)$, que es diferenciable en $\vec{a} = (a_1, a_2, \ldots, a_n) \in D$. Entonces, la diferencial de f, $df(\vec{a})$, es una transformación lineal con dominio $\mathbb{R}^n$ y valores en $\mathbb{R}$; por tanto, basta conocer una representación matricial de $df(\vec{a})$ para saber los valores de esta transformación en cualquier $\vec{u} = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$. Sean $\vec{e}_i$, $i = 1, 2, \ldots, n$, los vectores de la base canónica de $\mathbb{R}^n$, la representación matricial de $df(\vec{a})$ respecto a esta base está dada por la matriz $1 \times n$

$$\begin{bmatrix} df(\vec{a})\vec{e}_1 & df(\vec{a})\vec{e}_2 & \cdots & df(\vec{a})\vec{e}_n \end{bmatrix};$$

pero

$$\begin{aligned} df(\vec{a})\vec{e}_i &= \partial_{\vec{e}_i} f(\vec{a}) \\ &= \lim_{h\to 0} \frac{f(\vec{a}+h\vec{e}_i) - f(\vec{a})}{h} \\ &= \lim_{h\to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n)}{h} \end{aligned}$$

Este último límite, cuando existe, se llama la **derivada parcial de f en el punto $\vec{a}$ respecto a la variable** x_i y se denota por

$$\frac{\partial f}{\partial x_i}(\vec{a})$$

o bien

$$f_{x_i}(\vec{a})$$

y representa la razón de cambio que tiene esta función en el punto $\vec{a}$ respecto a la variable x_i cuando todas las demás variables se matienen constantes. Por tanto,

$$\left[\ df(\vec{a})\vec{e}_1 \quad df(\vec{a})\vec{e}_2 \quad \cdots \quad df(\vec{a})\vec{e}_n\ \right] = \left[\ \frac{\partial f}{\partial x_1}(\vec{a}) \quad \frac{\partial f}{\partial x_2}(\vec{a}) \quad \cdots \quad \frac{\partial f}{\partial x_n}(\vec{a})\ \right]$$

y, entonces,

$$df(\vec{a})\vec{u} = \left[\ \frac{\partial f}{\partial x_1}(\vec{a}) \quad \frac{\partial f}{\partial x_2}(\vec{a}) \quad \cdots \quad \frac{\partial f}{\partial x_n}(\vec{a})\ \right]\vec{u} \quad \forall \vec{u} \in \mathbb{R}^n.$$

A la representación matricial respecto a la base canónica de la diferencial de f en $\vec{a}$,

$$[df(\vec{a})] = \left[\ \frac{\partial f}{\partial x_1}(\vec{a}) \quad \frac{\partial f}{\partial x_2}(\vec{a}) \quad \cdots \quad \frac{\partial f}{\partial x_n}(\vec{a})\ \right],$$

se le llama **matriz jacobiana** del funcional f en el punto $\vec{a}$. Por ello, si la función f es diferenciable en $\vec{a}$, todas las derivadas parciales, $\frac{\partial f}{\partial x_i}(\vec{a})$, existen y

$$df(\vec{a})\vec{u} = [df(\vec{a})]\,\vec{u}.$$

Condiciones necesarias para diferenciabilidad de funcionales en $\mathbb{R}^n$ y cálculo de derivadas parciales

Por tanto, una condición necesaria para que f sea diferenciable en $\vec{a}$ es que existan todas las derivadas parciales de f en $\vec{a}$. Sea $i \in \{1, 2, \ldots, n\}$ un índice fijo, se define la función de variable real

$$g(x) = f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n)$$

entonces,

$$\lim_{h\to 0} \frac{g(a_i + h) - g(a)}{h} = \lim_{h\to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n)}{h}.$$

Luego

$$\frac{\partial f}{\partial x_i}(\vec{a}) \text{ existe } \Leftrightarrow g'(a_i) \text{ existe}$$

y entonces

$$\frac{\partial f}{\partial x_i}(\vec{a}) = g'(a_i).$$

► **Ejemplo 6.68** Sean

$$f(x_1, x_2, x_3, x_4) = x_1 x_2^2 - \cos(x_1 x_2 + x_3 x_4),$$

$\vec{a} = (x_1, x_2, x_3, x_4)$, e $i = 2$; entonces

$$g(x) = x_1 x^2 - \cos(x_1 x + x_3 x_4) \qquad \text{y}$$
$$g'(x) = 2x_1 x + x_1 \operatorname{sen}(x_1 x + x_3 x_4)$$

Claramente g es derivable en todo punto, por tanto,

$$\frac{\partial f}{\partial x_2}(x_1, x_2, x_3, x_4) = 2x_1 x_2 + x_1 \operatorname{sen}(x_1 x_2 + x_3 x_4).$$

Análogamente

$$\frac{\partial f}{\partial x_1}(x_1,x_2,x_3,x_4) = x_2^2 + x_2\operatorname{sen}(x_1x_2 + x_3x_4),$$

$$\frac{\partial f}{\partial x_3}(x_1,x_2,x_3,x_4) = x_4\operatorname{sen}(x_1x_2 + x_3x_4) \qquad \text{y}$$

$$\frac{\partial f}{\partial x_4}(x_1,x_2,x_3,x_4) = x_3\operatorname{sen}(x_1x_2 + x_3x_4).◄$$

Entonces, para determinar si una función tiene derivada parcial respecto a x_i, se mantienen las restantes variables fijas (como si fueran constantes) y si la función resultante, $g(x_i)$, es derivable, existe la derivada parcial respecto a x_i y además, $g'(x_i)$ es esa derivada parcial.

► **Ejemplo 6.69** Si

$$f(x,y) = xe^{xy^2} + \frac{x}{y},$$

entonces

$$\frac{\partial f}{\partial x}(x,y) = \frac{\partial}{\partial x}\left(xe^{xy^2} + \frac{x}{y}\right) = (1 + xy^2)e^{xy^2} + \frac{1}{y}$$

y

$$\frac{\partial f}{\partial y}(x,y) = \frac{\partial}{\partial y}\left(xe^{xy^2} + \frac{x}{y}\right) = 2x^2ye^{xy^2} - \frac{x}{y^2}.◄$$

Las derivadas parciales tienen una interpretación geométrica sumamente interesante en el caso de que el funcional dependa de dos variables únicamente. Supongamos que las derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$ existen en un punto (a,b). Sean

$$g_1(x) = f(x,b) \text{ y}$$

$$g_2(y) = f(a,y),$$

entonces,

$$g_1'(a) = \frac{\partial f}{\partial x}(a,b) \text{ y } g'(b) = \frac{\partial f}{\partial y}(a,b).$$

Si consideramos el plano $y = b$, éste corta la gráfica de la función $z = f(x,y)$ en la curva $z = g_1(x)$; luego $g_1'(a) = \frac{\partial f}{\partial x}(a,b)$ es la pendiente de la recta tangente a esta curva en el punto $(a, g_1(a))$, como se ilustra en la figura 6-31. De manera análoga el plano $x = a$ corta la gráfica de la función $z = f(x,y)$ en la curva $z = g_2(y)$; así que $g_2'(b) = \frac{\partial f}{\partial y}(a,b)$ es la pendiente de la recta tangente a esa curva en el punto $(b, g_2(b))$, como se muestra en la figura 6-32.

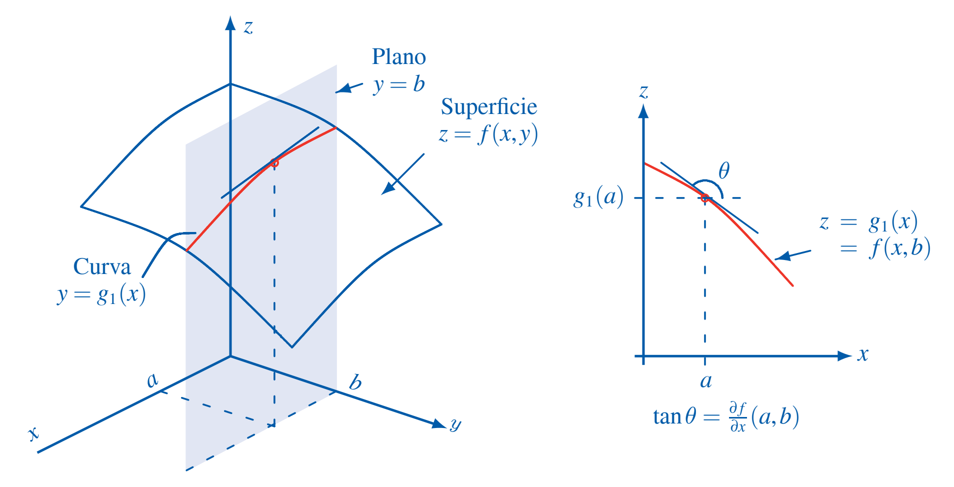


Figura 6-31 •

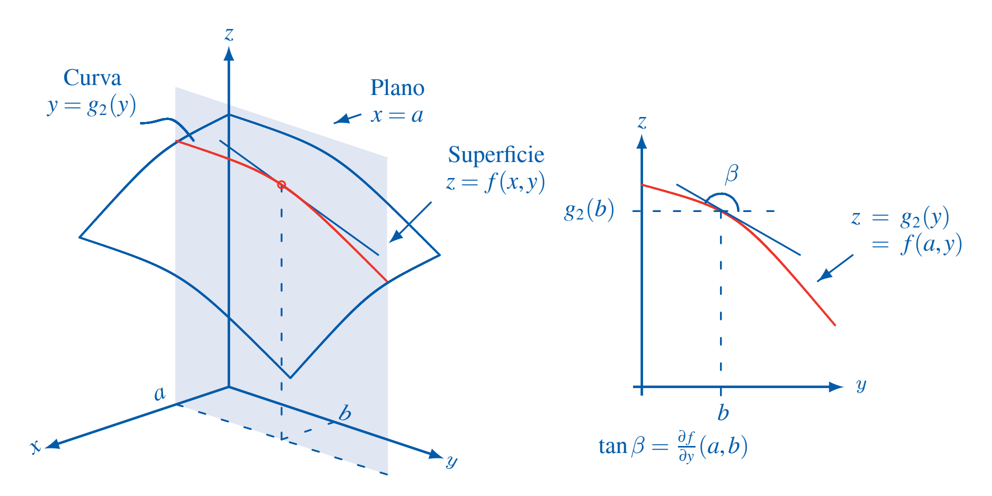


Figura 6-32 •

Condiciones suficientes de diferenciabilidad

Es posible que un funcional $f : D \subset \mathbb{R}^n \to \mathbb{R}$ tenga todas sus derivadas parciales en un punto $\vec{a}$ y no sea diferenciable en éste; es más, se pueden dar ejemplos de funciones para las cuales existe derivada direccional en la dirección de cualquier vector en un punto dado, pero que no son diferenciables en ese punto. Afortunadamente, para el caso de funcionales en $\mathbb{R}^n$ existen condiciones muy sencillas de verificar, con las derivadas parciales, que son suficientes para la diferenciabilidad del funcional. Antes de ello necesitamos recordar un teorema que es piedra angular en las matemáticas. En la figura 6-33 se tiene la interpretación geométrica de este importante teorema:

Teorema 6.17 **(Teorema del valor medio)** *Sea f una función de una variable real con valores reales. Si f es continua en todos los puntos del intervalo $[\alpha,\beta]$ y derivable en todos los puntos del intervalo (α,β), entonces existe $c \in (\alpha,\beta)$ tal que*

$$f'(c) = \frac{f(\beta) - f(\alpha)}{\beta - \alpha}.$$

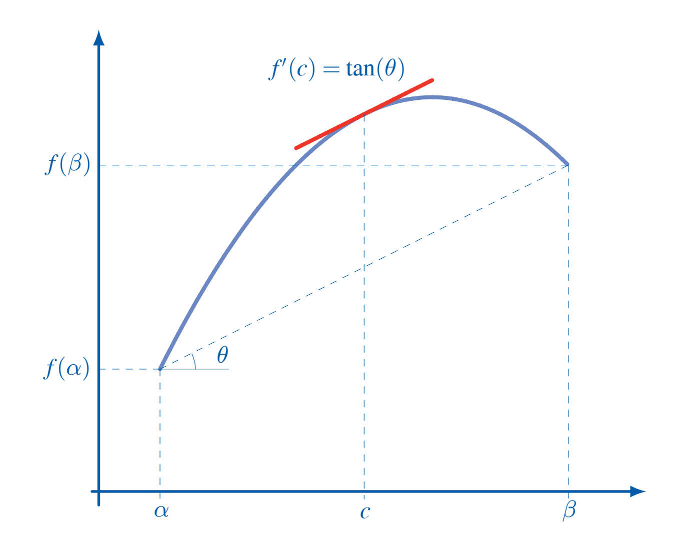


Figura 6-33 •

Ahora supongamos que $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ es un funcional y existe $r > 0$ tal que las derivadas parciales $\partial f/\partial x, \partial f/\partial y$ existen en todo punto de $B((a,b),r)$ y son continuas en (a,b). Sea la función lineal definida, para cada $\vec{u} = (u_1,u_2)$, por

$$L(u_1,u_2) = \left[\ \frac{\partial f}{\partial x}(a,b) \quad \frac{\partial f}{\partial y}(a,b)\ \right]\vec{u};$$

esto es

$$L(u_1,u_2) = \frac{\partial f}{\partial x}(a,b)u_1 + \frac{\partial f}{\partial y}(a,b)u_2.$$

L es continua porque $\mathbb{R}^2$ tiene dimensión finita. Vamos a probar que

$$\lim_{(h_1,h_2)\to(0,0)} \frac{f(a+h_1,b+h_2) - f(a,b) - L(h_1,h_2)}{\|(h_1,h_2)\|} = 0\,.$$

Tenemos que

$$f(a+h_1,b+h_2) - f(a,b) = f(a+h_1,b+h_2) - f(a,b+h_2) + f(a,b+h_2) - f(a,b)$$

y aplicando el teorema del valor medio, existen c_1, entre a y $a+h_1$; c_2, entre b y $b+h_2$, tales que

$$\begin{aligned} f(a+h_1,b+h_2) - f(a,b+h_2) &= \frac{\partial f}{\partial x}(c_1,b+h_2)h_1 \text{ y} \\ f(a,b+h_2) - f(a,b) &= \frac{\partial f}{\partial y}(a,c_2)h_2. \end{aligned}$$

Por tanto,

$$\begin{aligned} \frac{|f(a+h_1,b+h_2) - f(a,b) - L(h_1,h_2)|}{\|(h_1,h_2)\|} &= \frac{1}{\|(h_1,h_2)\|}\left|\left(\frac{\partial f}{\partial x}(c_1,b+h_2) - \frac{\partial f}{\partial x}(a,b)\right)h_1\right. \\ &\qquad \left. + \left(\frac{\partial f}{\partial y}(a,c_2) - \frac{\partial f}{\partial y}(a,b)\right)h_2\right| \end{aligned}$$

$$\leq \frac{1}{\|(h_1,h_2)\|}\left[\|(h_1,h_2)\|\left|\frac{\partial f}{\partial x}(c_1,b+h_2)-\frac{\partial f}{\partial x}(a,b)\right|+\|(h_1,h_2)\|\left|\frac{\partial f}{\partial y}(a,c_2)-\frac{\partial f}{\partial y}(a,b)\right|\right]$$
$$=\left|\frac{\partial f}{\partial x}(c_1,b+h_2)-\frac{\partial f}{\partial x}(a,b)\right|+\left|\frac{\partial f}{\partial y}(a,c_2)-\frac{\partial f}{\partial y}(a,b)\right|$$

y ya que ambas parciales son continuas y $c_1 \to a$, $c_2 \to b$ cuando $(h_1,h_2) \to (0,0)$,

$$\lim_{(h_1,h_2)\to(0,0)}\left|\frac{\partial f}{\partial x}(c_1,b+h_2)-\frac{\partial f}{\partial x}(a,b)\right|=0,$$
$$\lim_{(h_1,h_2)\to(0,0)}\left|\frac{\partial f}{\partial y}(a,c_2)-\frac{\partial f}{\partial y}(a,b)\right|=0.$$

De donde

$$\lim_{(h_1,h_2)\to(0,0)}\frac{|f(a+h_1,b+h_2)-f(a,b)-L(h_1,h_2)|}{\|(h_1,h_2)\|}=0.$$

Por tanto, f es diferenciable en (a,b). La demostración del caso general se hace por analogía y se deja de ejercicio al lector. Resumimos la precedente información en el siguiente teorema.

Teorema 6.18 *Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ un funcional. Se supone que $\vec{a} \in D$ y existe $r > 0$ tal que:*

1. *Las derivadas parciales $\frac{\partial f}{\partial x_i}$, $i = 1,2,\ldots,n$, existen en todo punto de $B(\vec{a},r)$.*
2. *Todas las derivadas parciales $\frac{\partial f}{\partial x_i}$ son continuas en $\vec{a}$.*

Entonces, f es diferenciable en $\vec{a}$.

► **Ejemplo 6.70** Sean $f(x,y) = x\sqrt{y}$ y $(a,b) = (2,3)$. Entonces,

$$\frac{\partial f}{\partial x}(x,y)=\sqrt{y} \quad \text{y} \quad \frac{\partial f}{\partial y}(x,y)=\frac{x}{2\sqrt{y}}\,.$$

Ambas derivadas parciales existen y son continuas en todo punto (x,y) con $x \in \mathbb{R}$ y $y \in (0,\infty)$ (el lector debe utilizar los resultados vistos en el apartado de límites y continuidad para justificar esta afirmación). Para este caso particular podemos tomar $r = 2$ y las derivadas parciales existen en todo punto de $B((2,3),r)$ y son continuas en $(2,3)$. Por tanto, f es diferenciable en (2,3) y

$$\begin{aligned} df(2,3)(u_1,u_2) &= [df(2,3)]\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &= \begin{bmatrix} f_x(2,3) & f_y(2,3) \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{3} & \sqrt{3}/3 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &= \sqrt{3}(u_1+(1/3)u_2). \end{aligned}$$ ◄

► **Ejemplo 6.71** Sea $f(x,y,z) = xy^2z - xye^z$. Entonces

$$f_x(x,y,z) = y^2z - ye^z,$$
$$f_y(x,y,z) = 2xyz - xe^z,$$
$$f_z(x,y,z) = xy^2 - xye^z.$$

Por tanto, las derivadas parciales existen y son continuas[20] en todo punto de $\mathbb{R}^3$; luego f es diferenciable en todo punto. Además,

$$df(x,y,z)\vec{u} = \begin{bmatrix} y^2z - ye^z & 2xyz - xe^z & xy^2 - xye^z \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

para todo $(x,y,z) \in \mathbb{R}^3$ y para todo $\vec{u} = (u_1,u_2,u_3) \in \mathbb{R}^3$.◄

Condiciones suficientes de diferenciabilidad y diferencial de funciones $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$

Sean $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ una función y $f_i : D \to \mathbb{R}$, $i = 1,\ldots,n$, sus funciones coordenadas; esto es,

$$f(\vec{u}) = (f_1(\vec{u}), f_2(\vec{u}), \ldots, f_m(\vec{u})) \;\; \forall \vec{u} \in D.$$

Supongamos que f es diferenciable en $\vec{a} \in D$ y sean $L_1,\ldots,L_m$ las funciones coordenadas de $df(\vec{a})$, entonces cada L_i también es lineal (cfr. ejercicio propuesto 62 del cap. 5). Sean $\vec{e}_i$ los vectores de la base canónica de $\mathbb{R}^n$, entonces

$$\vec{0}_{\mathbb{R}^m} = \lim_{\vec{h}\to\vec{0}_{\mathbb{R}^n}} \frac{f(\vec{a}+\vec{h}) - f(\vec{a}) - df(\vec{a})\vec{h}}{\|\vec{h}\|} = \lim_{\vec{h}\to\vec{0}_{\mathbb{R}^n}} \sum_{i=1}^{m} \frac{f_i(\vec{a}+\vec{h}) - f_i(\vec{a}) - L_i\vec{h}}{\|\vec{h}\|}\vec{e}_i$$

implica

$$\lim_{\vec{h}\to\vec{0}_{\mathbb{R}^n}} \frac{f_i(\vec{a}+\vec{h}) - f_i(\vec{a}) - L_i\vec{h}}{\|\vec{h}\|} = \vec{0}_{\mathbb{R}^m}$$

para todo $i = 1,\ldots,m$. Es decir, la funciones coordenadas son también diferenciables en $\vec{a}$ y $L_i = df_i(a)$ para todo i. Inversamente, si todas las funciones coordenadas son diferenciables en $\vec{a}$, $L = (df_1(\vec{a}), df_2(\vec{a}), \ldots, df_m(\vec{a}))$ es lineal y

$$\vec{0}_{\mathbb{R}^m} = \lim_{\vec{h}\to\vec{0}_{\mathbb{R}^n}} \sum_{i=1}^{m} \frac{f_i(\vec{a}+\vec{h}) - f_i(\vec{a}) - df_i(a)\vec{h}}{\|\vec{h}\|}\vec{e}_i = \lim_{\vec{h}\to\vec{0}_{\mathbb{R}^n}} \frac{f(\vec{a}+\vec{h}) - f(\vec{a}) - L\vec{h}}{\|\vec{h}\|}$$

y, por tanto, f es diferenciable[21] en $\vec{a}$ y $df(\vec{a})\vec{u} = (df_1(\vec{a})\vec{u}, df_2(\vec{a})\vec{u}, \ldots, df_m(\vec{a})\vec{u})$. Esto es, si $\vec{u} = (u_1,u_2,\ldots,u_n)$

[20] Recuerde que el propósito del apartado de continuidad y límites fue el de poder determinar por simple inspección la continuidad de funciones elementales que se presentan con mucha frecuencia; usted debe ser capaz de justificar las afirmaciones de continuidad que presentamos en este ejemplo y los siguientes, aunque las argumentaciones no se escriban.

[21] Observe que no necesitamos demostrar la continuidad de la transformación lineal que se propone como diferencial porque el dominio es $\mathbb{R}^n$, que tiene dimensión finita y, por tanto, toda transformación lineal en él es continua.

$$df(\vec{a})\vec{u} = \left(\sum_{j=1}^{n}\frac{\partial f_1}{\partial x_i}(\vec{a})u_i, \sum_{j=1}^{n}\frac{\partial f_2}{\partial x_i}(\vec{a})u_i, \ldots, \sum_{j=1}^{n}\frac{\partial f_m}{\partial x_i}(\vec{a})u_i\right)$$

$$= \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}) & \frac{\partial f_1}{\partial x_2}(\vec{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{a}) \\ \frac{\partial f_2}{\partial x_1}(\vec{a}) & \frac{\partial f_2}{\partial x_2}(\vec{a}) & \cdots & \frac{\partial f_2}{\partial x_n}(\vec{a}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{a}) & \frac{\partial f_m}{\partial x_2}(\vec{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\vec{a}) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}.$$

La matriz

$$[df(\vec{a})] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}) & \frac{\partial f_1}{\partial x_2}(\vec{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{a}) \\ \frac{\partial f_2}{\partial x_1}(\vec{a}) & \frac{\partial f_2}{\partial x_2}(\vec{a}) & \cdots & \frac{\partial f_2}{\partial x_n}(\vec{a}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{a}) & \frac{\partial f_m}{\partial x_2}(\vec{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\vec{a}) \end{bmatrix} \tag{6.66}$$

se llama **matriz jacobiana** de la función f en el punto $\vec{a}$ y no es más que la representación matricial de la diferencial de f en ese punto, relativa a las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$. Empleando el teorema 6.18 tenemos como consecuencia inmediata de la discusión precedente, el siguiente resultado que establece condiciones para que una función sea diferenciable cuando su dominio y contradominio están contenidos en espacios $\mathbb{R}^k$.

Teorema 6.19 *Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ una función con funciones coordenadas f_j, $j = 1,\ldots,m$, y $\vec{a} \in D$. Si existe $r > 0$, tal que para cada $j = 1,\ldots,m$ las derivadas parciales $\partial f_j/\partial x_i$, $i = 1,\ldots,n$, existen en todo punto de $B(\vec{a},r)$ y son continuas en $\vec{a}$, entonces f es diferenciable en $\vec{a}$. Además,*

$$df(\vec{a}) = [df(\vec{a})]\vec{u}$$

donde $[df(\vec{a})]$ es la matriz jacobiana de f en $\vec{a}$ dada por (6.66).

► **Ejemplo 6.72** Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ definida por

$$f(x,y,z) = (xyz^2, xye^{yz^2}).$$

Determinar los puntos donde f es diferenciable, calcular la matriz jacobiana en cualquier punto (x,y,z) donde f es diferenciable y calcular $df(1,0,1)\vec{u}$.◄

Solución En este caso las funciones coordenadas de f son $f_1(x,y,z) = xyz^2$ y $f_2(x,y,z) = xye^{yz^2}$. Entonces,

$$\begin{aligned} \frac{\partial f_1}{\partial x}(x,y,z) &= \frac{\partial}{\partial x}(xyz^2) = yz^2, \\ \frac{\partial f_1}{\partial y}(x,y,z) &= \frac{\partial}{\partial y}(xyz^2) = xz^2, \\ \frac{\partial f_1}{\partial z}(x,y,z) &= \frac{\partial}{\partial z}(xyz^2) = 2xyz; \end{aligned}$$

$$\frac{\partial f_2}{\partial x}(x,y,z) = \frac{\partial}{\partial x}(xye^{yz^2}) = ye^{yz^2},$$

$$\frac{\partial f_2}{\partial y}(x,y,z) = \frac{\partial}{\partial y}(xye^{yz^2}) = x(1+yz^2)e^{yz^2},$$

$$\frac{\partial f_2}{\partial z}(x,y,z) = \frac{\partial}{\partial z}(xye^{yz^2}) = 2xy^2ze^{yz^2}.$$

Como las parciales de f_1 y f_2 respecto a x, y, z existen en todo punto y son continuas en todo punto, entonces f es diferenciable en todo $(x,y,z) \in \mathbb{R}^3$. La matriz jacobiana de f en un punto (x,y,z) está dada por

$$[df(x,y,z)] = \begin{bmatrix} yz^2 & xz^2 & 2xyz \\ ye^{yz^2} & x(1+yz^2)e^{yz^2} & 2xy^2ze^{yz^2} \end{bmatrix}$$

y, por tanto, si $(x,y,z) = (1,0,1)$ y $\vec{u} = (u_1,u_2,u_3)$, entonces

$$df(1,0,1)\vec{u} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = (u_2,u_2). \quad ✔$$

6.8.4 Extremos locales de funcionales

Definición 6.12 *Sean* $\mathbf{E}$ *un espacio vectorial normado,* $J : D \subset \mathbf{E} \to \mathbb{R}$ *y* $f_0 \in D$.

1. *Se dice que* J *tiene (o alcanza) un* ***máximo local (relativo)*** *en* f_0, *si existe* $r > 0$ *tal que*

$$J(f) \leq J(f_0)$$

para todo $f \in B(f_0,r)$. *En tal caso al valor real* $J(f_0)$ *se le llama* ***máximo local (relativo)*** *del funcional* J.

2. *Se dice que* J *tiene (o alcanza) un* ***mínimo local (relativo)*** *en* f_0, *si existe* $r > 0$ *tal que*

$$J(f_0) \leq J(f)$$

para todo $f \in B(f_0,r)$. *En tal caso al valor real* $J(f_0)$ *se le llama* ***mínimo local (relativo)*** *del funcional* J.

3. *Los máximos y mínimos locales (relativos) de un funcional se llaman* ***extremos locales (relativos)***.

En el caso de un funcional de dos variables reales, $z = f(x,y)$, los máximos relativos corresponden en la gráfica de la función a los puntos más altos de "montañas" y los mínimos relativos a los puntos más bajos de "valles" o "sumideros", como se ilustra en la figura 6-34.

Supongamos que $J : D \subset \mathbf{E} \to \mathbb{R}$ tiene un máximo local en f_0 y que J es diferenciable en f_0. Sea $u \in \mathbf{E}$, entonces, para t suficientemente pequeño, $J(f_0+tu) - J(f_0) \leq 0$, por tanto,

$$\begin{aligned} dJ(f_0)\vec{u} &= \partial_u J(f_0) \\ &= \lim_{t \to 0^+} \frac{J(f_0+tu) - J(f_0)}{t} \leq 0 \end{aligned}$$

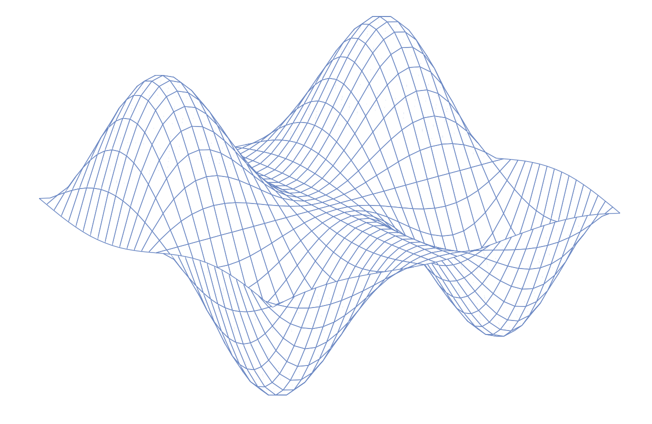


Figura 6-34 •

y

$$\begin{aligned} dJ(f_0)\vec{u} &= \partial_u J(f_0) \\ &= \lim_{t\to 0^-} \frac{J(f_0+tu)-J(f_0)}{t} \geq 0 \end{aligned}$$

De donde se deduce

$$dJ(f_0)\vec{u} = 0\,.$$

El caso en el que J alcanza un mínimo es análogo. Hemos probado así el siguiente teorema.

Teorema 6.20 *Sea J un funcional diferenciable en f_0. Si el funcional J alcanza un máximo o mínimo relativo en f_0, entonces*

$$dJ(f_0)u = 0 \ \ \forall u.$$

Es decir, $dJ(f_0)$ es la función constante cero.

Definición 6.13 *Sea $J : D \subset \mathbf{E} \to \mathbb{R}$ un funcional. $f_0 \in D$ es un **punto crítico** de J, si el funcional es diferenciable en f_0 y la diferencial de J en f_0 es la función constante cero: $dJ(f)u = 0\, \forall\, u \in \mathbf{E}$.*

❍ **Nota 6.6** De manera análoga al caso de funciones $J : U \subset \mathbb{R} \to \mathbb{R}$, el teorema anterior sólo da condiciones necesarias para que J tenga un extremo relativo en f_0; es decir, puede ser que $dJ\,(f_0) = 0$, pero $J\,(f_0)$ no sea un extremo relativo. Por ejemplo, la función $J\,(x) = x^3$ satisface $J'\,(x) = 0$ en $x = 0$, pero $J\,(0) = 0$ no es un extremo relativo.

► **Ejemplo 6.73** Sea $\mathbf{E} = C\,[a,b]$, y J: $\mathbf{E} \to \mathbb{R}$ el funcional definido por

$$J\,(u) = \frac{1}{2}\int_a^b u^2$$

Entonces

$$\lim_{h\to 0} \frac{J\,(u+h) - J\,(u) - \displaystyle\int_a^b uh}{\|h\|_\infty} = \lim_{h\to 0} \frac{\frac{1}{2}\displaystyle\int_a^b (u+h)^2 - \frac{1}{2}\int_a^b u^2 - \int_a^b uh}{\|h\|_\infty} = \lim_{h\to 0} \frac{\frac{1}{2}\displaystyle\int_a^b h^2}{\|h\|_\infty} = 0$$

pues

$$0 \leq \lim_{h \to 0} \frac{\int_a^b h^2}{\|h\|_\infty} \leq \lim_{h \to 0} \frac{\int_a^b \|h\|_\infty^2}{\|h\|_\infty} = \lim_{h \to 0} \frac{(b-a)\|h\|_\infty^2}{\|h\|_\infty} = 0$$

Por otra parte, la aplicación $J'(u) : \mathbf{E} \to \mathbb{R}$ definida por

$$J'(u)h = \int_a^b uh \tag{6.67}$$

es claramente lineal y, además,

$$\begin{aligned} |J'(u)h| &= \left| \int_a^b uh \right| \\ &\leq \int_a^b |uh| \\ &\leq \int_a^b \|u\|_\infty \|h\|_\infty \\ &= (b-a)\|u\|_\infty \|h\|_\infty \\ &= M\|h\|_\infty \end{aligned}$$

lo cual implica que J es diferenciable en todo punto de $\mathbf{E}$ y $dJ(u) = J'(u)$, la cual está definida por (6.67).◀

▶ **Ejemplo 6.74** Encontrar los puntos críticos del funcional $J : C[a,b] \to \mathbb{R}$ definido por

$$J(u) = \frac{1}{2}\int_a^b u^2 \blacktriangleleft$$

Solución Por el ejemplo 6.73, J es diferenciable en todo punto de $C[a,b]$ y, además, $J'(u)h = \int_a^b uh$. Luego, si u es un punto crítico de J, se debe cumplir

$$\int_a^b uh = 0$$

para todo $h \in C[a,b]$. Esto implica que el producto interior de u con cualquier función h en $C[a,b]$ es cero y, por tanto, u debe ser la función constante cero en $[a,b]$. Así que este funcional sólo tiene un punto crítico: $u = 0$. ✔

Es claro que el funcional del ejemplo precedente tiene un mínimo absoluto en $u = 0$, el cual también es mínimo relativo. Sin embargo, puede suceder que un punto crítico no sea un extremo relativo como ya hemos mencionado antes.

▶ **Ejemplo 6.75** Sea $J : C[a,b] \to \mathbb{R}$ el funcional definido por $J(u) = \frac{1}{3}\int_a^b u^3$. No es difícil ver que J es diferenciable en todo punto de $C[a,b]$ y que además,[22] $J'(u)h = \int_a^b u^2 h$. De donde, de nuevo, el producto interior de u^2 con cualquier otra función h en $C[a,b]$ es cero y, por tanto, u debe ser la función constante cero; así J tiene un solo punto crítico: $u = 0$. Sin embargo, las funciones constantes $u_1(x) = \varepsilon$ y $u_2(x) = -\varepsilon$ son arbitrariamente cercanas a $u = 0$, tomando $\varepsilon > 0$ lo suficiente pequeño y $J(u_1) > J(u) > J(u_2)$; por lo que el punto crítico $u = 0$ no es un extremo relativo.◀

[22] Esto se deja al lector como ejercicio.

6.8.5 Extremos locales y valores propios de la matriz hessiana

En el caso de que f sea un funcional en $\mathbb{R}^n$, entonces $df(\vec{a})\vec{u} = 0$ para todo $\vec{u}$, si y sólo si la matriz jacobiana es la matriz cero; esto es, todas las derivadas parciales son nulas en $\vec{a}$.

Teorema 6.21 (Puntos críticos de un funcional en $\mathbb{R}^n$) *Si $f : D \subset \mathbb{R}^n \to \mathbb{R}$ es un funcional, entonces $\vec{a} \in D$ es punto crítico de f, si y sólo si*

$$\frac{\partial f}{\partial x_i}(\vec{a}) = 0$$

para todo $i = 1, \ldots, n$.

Condiciones suficientes para alcanzar valores extremos de funcionales en $\mathbb{R}^n$

Para funcionales con dominio en $\mathbb{R}^n$ existen condiciones que son suficientes para alcanzar extremos locales en puntos críticos. Antes de establecer estas condiciones necesitamos de algunos aspectos técnicos y operativos que presentamos a continuación.

Para fijar ideas primero, sea $z = f(x,y)$ un funcional que tiene derivadas parciales f_x y f_y. Entonces se definen las segundas derivadas parciales como:

- $$\frac{\partial f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \text{ y}$$
  $$\frac{\partial f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right).$$

- (Derivadas parciales mixtas)

  $$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \text{ y}$$
  $$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

- O con la notación equivalente:

  $$f_{xx} = (f_x)_x\,,$$
  $$f_{yy} = (f_y)_y\,,$$
  $$f_{xy} = (f_x)_y\,,$$
  $$f_{yx} = (f_y)_x\,.$$

siempre que éstas existan.

Se puede probar que si las derivadas parciales mixtas, f_{xy} y f_{yx}, existen en una bola abierta de centro (a,b) y radio r y son continuas en (a,b), entonces $f_{xy}(a,b) = f_{yx}(a,b)$. Es claro que lo precedente se puede generalizar a funcionales con dominio en $\mathbb{R}^n$:

- Segundas derivadas parciales

  $$f_{x_i x_i} = (f_{x_i})_{x_i}$$

(derivadas parciales mixtas) si $i \neq j$,

$$f_{x_i x_j} = (f_{x_i})_{x_j}$$

- Las **derivadas parciales mixtas son iguales en aquellos puntos donde poseen continuidad**

Definición 6.14 *Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ un funcional para el que todas las segundas derivadas parciales $f_{x_i x_j}$, $i = 1, \ldots n$, $j = 1, \ldots, n$, existen en la bola $B(\vec{a}, r)$ y son continuas en $\vec{a}$. La matriz simétrica*

$$H(\vec{a}) = \left[f_{x_i x_j} \right]$$

*$i = 1, \ldots, n$, $j = 1, \ldots n$, se llama **matriz hessiana** de f en $\vec{a}$.*

Teorema 6.22 **(Fórmula de Taylor de segundo orden para funcionales en $\mathbb{R}^n$)** *Sean $f : D \subset \mathbb{R}^n \to \mathbb{R}$ un funcional y $\vec{a} \in D$, tal que las segundas derivadas parciales $f_{x_i x_j}$, $i = 1, \ldots n$, $j = 1, \ldots, n$, existen y son continuas en todos los puntos de la bola $B(\vec{a}, r)$.*

1. *Entonces, para cada $\vec{h} \in \mathbb{R}^n$ con $\vec{a} + \vec{h} \in B(\vec{a}, r)$, existe $c \in (0, 1)$ tal que*

$$f(\vec{a} + \vec{h}) - f(\vec{a}) = df(\vec{a})\vec{h} + \frac{1}{2}\vec{h}^t H(\vec{a} + c\vec{h})\vec{h} \tag{6.68}$$

2. *Existe una función $E_2(\vec{a}, \vec{h})$ tal que*

$$f(\vec{a} + \vec{h}) - f(\vec{a}) = df(\vec{a})\vec{h} + \frac{1}{2}\vec{h}^t H(\vec{a})\vec{h} + \|\vec{h}\|^2 E_2(\vec{a}, \vec{h}) \tag{6.69}$$

   y

$$\lim_{\vec{h} \to \vec{0}_{\mathbb{R}^n}} E_2(\vec{a}, \vec{h}) = 0.$$

DEMOSTRACIÓN 1. Sea $\varphi : [-1, 1] \to \mathbb{R}$, la función definida por

$$\varphi(t) = f(\vec{a} + t\vec{h}).$$

Luego,

$$\varphi(1) - \varphi(0) = f(\vec{a} + \vec{h}) - f(\vec{a}).$$

Apliquemos el teorema de Taylor de segundo[23] orden para funciones reales de variable real a la función φ, entonces existe $c \in (0, 1)$ tal que

$$\varphi(1) - \varphi(0) = \varphi'(0) + \frac{1}{2}\varphi''(c).$$

Ahora bien, $\varphi = f \circ \rho$, donde $\rho(t) = \vec{a} + t\vec{h}$; es decir, $\varphi(t) = f(\rho(t))$. Entonces, si $\vec{a} = (a_1, \ldots, a_n)$ y $\vec{h} = (h_1, \ldots, h_n)$,

$$\rho(t) = (a_1 + th_1, \ldots, a_n + th_n)$$

[23] Si f es derivable con continuidad en todos los puntos de un intervalo abierto J, $x_0 \in J$ y $x_0 + h \in J$, entonces existe $c \in (0, 1)$ tal que $f(x_0 + h) = f(x_0) + f'(x_0)h + \frac{1}{2}f''(x_0)h^2$.

y, por tanto,

$$[d\rho(t)] = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}$$

Así, por la regla de cadena,

$$\begin{aligned}
[d\varphi(t)] &= [df(\rho(t))]\,[d\rho(t)] \\
&= \begin{bmatrix} f_{x_1}(\rho(t)) & f_{x_2}(\rho(t)) & \cdots & f_{x_n}(\rho(t)) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}
\end{aligned}$$

de donde

$$\varphi'(t) = \sum_{j=1}^{n} h_j f_{x_j}(\rho(t)).$$

Nuevamente al aplicar la regla de la cadena

$$\begin{aligned}
d\varphi'(t)\vec{h} &= d\left(\sum_{j=1}^{n} h_j f_{x_j}(\rho(t))\right)\vec{h} \\
&= \sum_{j=1}^{n} h_j d\left(f_{x_j}(\rho(t))\right)\vec{h} \\
&= \sum_{j=1}^{n} h_j \begin{bmatrix} f_{x_jx_1}(\rho(t)) & f_{x_jx_2}(\rho(t)) & \cdots & f_{x_jx_n}(\rho(t)) \end{bmatrix}\vec{h} \\
&= \sum_{j=1}^{n} h_j \left(\sum_{i=1}^{n} f_{x_jx_i}(\rho(t))h_i\right) \\
&= \sum_{j=1}^{n}\sum_{i=1}^{n} h_j f_{x_jx_i}(\rho(t))h_i \\
&= \begin{bmatrix} h_1 & h_2 & \cdots & h_n \end{bmatrix}^t H(\vec{a}+t\vec{h})\vec{h}.
\end{aligned}$$

Por tanto,

$$\begin{aligned}
f(\vec{a}+\vec{h}) - f(\vec{a}) &= \varphi(1) - \varphi(0) \\
&= df(\vec{a})\vec{h} + \frac{1}{2}\vec{h}^t H(a+c\vec{h})\vec{h}
\end{aligned}$$

y queda probado (6.68).

2. Sea

$$E_2(\vec{a},\vec{h}) = \begin{cases} \dfrac{\vec{h}^t}{2\|\vec{h}\|^2}\left(H(\vec{a}+c\vec{h}) - H(a)\right)\vec{h} & \text{si } \vec{h} \neq \vec{0}_{\mathbb{R}^n} \\ 0 & \text{si } \vec{h} = \vec{0}_{\mathbb{R}^n} \end{cases}$$

Entonces (6.68) se transforma en (6.69). Por tanto, si $\vec{h} \neq \vec{0}_{\mathbb{R}^n}$,

$$\begin{aligned}\|\vec{h}\|^2|E_2(\vec{a},\vec{h})| &= \frac{1}{2}\left(\vec{h}^t H(\vec{a}+c\vec{h})\vec{h} - \vec{h}^t H(a)\vec{h}\right)\\ &= \frac{1}{2}\left|\sum_{j=1}^{n}\sum_{i=1}^{n} h_j(f_{x_jx_i}(\vec{a}+c\vec{h}) - f_{x_jx_i}(\vec{a}))h_i\right|\\ &\leq \frac{1}{2}\|\vec{h}\|^2\sum_{j=1}^{n}\sum_{i=1}^{n}\left|(f_{x_jx_i}(\vec{a}+c\vec{h}) - f_{x_jx_i}(\vec{a}))\right|;\end{aligned}$$

de donde

$$\left|E_2(\vec{a},\vec{h})\right| \leq \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\left|(f_{x_jx_i}(\vec{a}+c\vec{h}) - f_{x_jx_i}(\vec{a}))\right|$$

y ya que las derivadas parciales son continuas,

$$\lim_{\vec{h}\to\vec{0}_{\mathbb{R}^n}} E_2(\vec{a},\vec{h}) = 0. \quad \blacksquare$$

Definición 6.15 *Sea A una matriz cuadrada de orden n con componentes reales y simétrica.*

1. *A es una matriz definida positiva, si*

$$\vec{x}^t A\vec{x} > 0$$

para todo $\vec{x} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$.

2. *A es una matriz definida negativa, si*

$$\vec{x}^t A\vec{x} < 0$$

para todo $\vec{x} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$.

Teorema 6.23 *Sea A una matriz simétrica real. Entonces,*

1. *A es definida positiva, si y sólo si todos sus valores propios son positivos.*
2. *A es definida negativa, si y sólo si todos sus valores propios son negativos.*

DEMOSTRACIÓN ■ Dado que A es una matriz con componentes reales y simétrica, todos los valores propios de A son números reales (cfr. teorema 5.39, pág. 494).

1. Supongamos que A es definida positiva; sea λ un valor propio de A con vector propio correspondiente $\vec{u}$. Entonces,

$$0 < \vec{u}^t A\vec{u} = \vec{u}^t \lambda\vec{u} = \lambda\,\|\vec{u}\|^2$$

De donde se desprende que $\lambda > 0$. Ahora supongamos que A tiene sus valores propios λ_i, $i = 1,\ldots,n$, positivos. Por el teorema 5.41, el espacio $\mathbb{R}^n$ tiene una base ortonormal de vectores propios $\vec{u}_i$ correspondientes a cada λ_i. Sean $\vec{u} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$ un vector cualquiera y escalares α_i tales que $\vec{u} = \sum_{i=1}^{n}\alpha_i\vec{u}_i$. Entonces, ya que

$$u_j^t \vec{u}_i = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

se tiene

$$\begin{aligned}
\vec{u}^{\,t} A \vec{u} &= \left(\sum_{j=1}^{n} \alpha_j \vec{u}_j^{\,t} \right) \left(\sum_{i=1}^{n} \alpha_i (A\vec{u}_i) \right) \\
&= \left(\sum_{j=1}^{n} \alpha_j \vec{u}_j^{\,t} \right) \left(\sum_{i=1}^{n} \alpha_i \lambda_i \vec{u}_i \right) \\
&= \left(\sum_{j=1}^{n} \sum_{i=1}^{n} \lambda_i \alpha_j \alpha_i \vec{u}_j^{\,t} \vec{u}_i \right) \\
&= \sum_{i=1}^{n} \alpha_i^2 \lambda_i \|\vec{u}_i\|^2 > 0
\end{aligned}$$

porque los λ_i son positivos, las normas $\|\vec{u}_i\|^2 > 0$ y al menos un α_i debe ser distinto a cero pues $\vec{u} \neq \vec{0}_{\mathbb{R}^n}$.

2. La demostración es completamente análoga al primer inciso. ■

Supongamos ahora que $\vec{a}$ es un punto crítico de f, entonces (6.69) se reduce a

$$f(a+\vec{h}) - f(\vec{a}) = \frac{1}{2}\vec{h}^{\,t} H(\vec{a})\vec{h} + \|\vec{h}\|^2 E_2(\vec{a},\vec{h}).$$

Puesto que $E_2(\vec{a},\vec{h}) \to 0$ cuando $\vec{h}$ tiende a cero, es de esperar que $f(a+\vec{h}) - f(\vec{a})$ y $\frac{1}{2}\vec{h}^{\,t} H(\vec{a})\vec{h}$ tengan el mismo signo. Así, si la matriz hessiana $H(\vec{a})$ es definida positiva, entonces $f(\vec{a}+\vec{h}) - f(\vec{a})$ sería mayor que cero para todo $\vec{h}$ suficientemente pequeño y entonces, $f(\vec{a})$ sería un mínimo relativo; y si $H(\vec{a})$ es definida negativa, entonces $f(\vec{a}+\vec{h}) - f(\vec{a})$ sería menor que cero para todo $\vec{h}$ lo bastante pequeño y, por tanto, $f(\vec{a})$ sería un máximo relativo.

Efectivamente, si la matriz hessiana $H(\vec{a})$ es definida positiva sean λ_i, $i = 1, \ldots, n$, sus valores propios y λ el menor de ellos. Entonces, si $0 < t < \lambda$, $\lambda_i - t > 0$ para todo i; así que la matriz simétrica

$$H(\vec{a}) - tI_n$$

es definida positiva porque sus valores propios son los $\lambda_i - t$. Luego,

$$\vec{h}^{\,t}(H(\vec{a}) - tI_n)\vec{h} > 0$$

para todo $\vec{h} \neq \vec{0}_{\mathbb{R}^n}$. Por tanto,

$$\vec{h}^{\,t} H(\vec{a})\vec{h} > t\|\vec{h}\|^2$$

para todo $0 < t < \lambda$. En particular para $t = \frac{1}{2}\lambda$; esto es

$$\vec{h}^{\,t} H(\vec{a})\vec{h} > \frac{1}{2}\lambda\|\vec{h}\|^2.$$

Dado que $E_2(\vec{a},\vec{h}) \to 0$ cuando $\vec{h} \to \vec{0}_{\mathbb{R}^n}$, existe $\delta > 0$ tal que

$$\left| E_2(\vec{a},\vec{h}) \right| < \frac{1}{4}\lambda$$

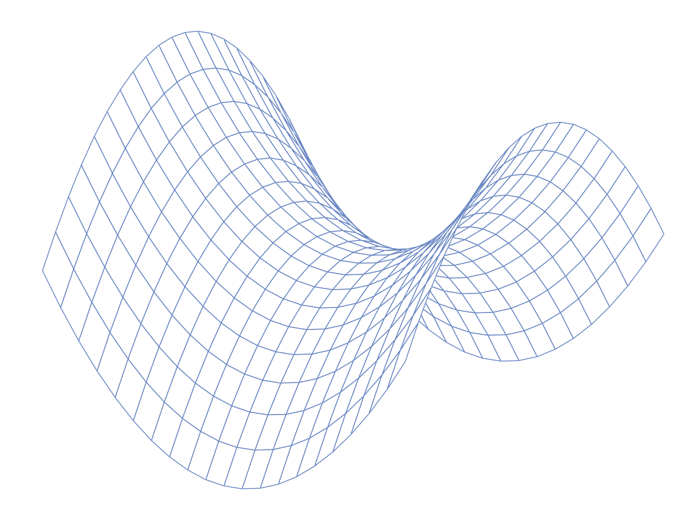


Figura 6-35 •

si $0 < \|\vec{h}\| < \delta$. Para todos estos valores se cumple entonces

$$\|\vec{h}\|^2 \left|E_2(\vec{a},\vec{h})\right| < \frac{1}{2}\frac{\|\vec{h}\|^2}{2}\lambda < \frac{1}{2}\vec{h}^t H(\vec{a})\vec{h}\,.$$

De donde se desprende

$$\begin{aligned} f(\vec{a}+\vec{h}) - f(\vec{a}) &= \frac{1}{2}\vec{h}^t H(\vec{a})\vec{h} + \|\vec{h}\|^2 E_2(\vec{a},\vec{h}) \\ &> \frac{1}{2}\vec{h}^t H(\vec{a})\vec{h} - \|\vec{h}\|^2 E_2(\vec{a},\vec{h}) \\ &> 0 \end{aligned}$$

para $0 < \|\vec{h}\| < \delta$. Luego $f(\vec{a})$ es un mínimo relativo. El caso en que $H(\vec{a})$ es definida negativa es completamente análogo. También se puede demostrar de manera similar que si $H(\vec{a})$ tiene dos valores propios λ_1 y λ_2, con $\lambda_1 < 0 < \lambda_2$, dado cualquier $\delta > 0$ existen $\vec{h}_i$, $i = 1, 2$, con $0 < \|\vec{h}_i\| < \delta$, tales que $f(\vec{a}+\vec{h}_1) - f(\vec{a}) > 0$ y $f(\vec{a}+\vec{h}_2) - f(\vec{a}) < 0$; lo cual significa que en $\vec{a}$ no se puede alcanzar ni máximo ni mínimo relativo. Se dice entonces, que en el punto $\vec{a}$ se tiene un **punto de silla**, o, punto silla, porque en el caso de una función de dos variables la gráfica, cerca de uno de estos puntos, se parece a una silla de montar, como se ilustra en la figura 6-35. Resumimos este resultado en el siguiente teorema.

Teorema 6.24 *Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ un funcional. Se supone que f tiene segundas derivadas parciales continuas en todos los puntos de una bola abierta que contiene al punto crítico $\vec{a}$ de f. Sea $H(\vec{a})$ la matriz hessiana de f en $\vec{a}$.*

1. *Si $H(\vec{a})$ tiene todos sus valores propios positivos, entonces f tiene un mínimo relativo en $\vec{a}$.*
2. *Si $H(\vec{a})$ tiene todos sus valores propios negativos, entonces f tiene un máximo relativo en $\vec{a}$.*
3. *Si $H(\vec{a})$ tiene valores propios positivos y negativos, entonces f tiene un punto de silla en $\vec{a}$.*

► **Ejemplo 6.76** Sea $f : \mathbb{R}^3 \to \mathbb{R}$ el funcional definido por

$$f(x,y,z) = \frac{1}{3}x^3 + \frac{1}{2}y^2 + z^2 - xy - 2z.$$

Encontrar los extremos relativos de f.◄

Solución Los puntos críticos de f son, por el teorema 6.21, las soluciones del sistema de ecuaciones

$$\begin{aligned} f_x &= 0 \\ f_y &= 0 \\ f_z &= 0 \end{aligned}$$

Entonces, ya que

$$\begin{aligned} f_x(x,y,z) &= x^2 - y \\ f_y(x,y,z) &= y - x \\ f_z(x,y,z) &= 2z - 2, \end{aligned}$$

los puntos críticos son $(0,0,1)$ y $(1,1,1)$. La matriz hessiana en cualquier punto (x,y,z) está dada por

$$\begin{aligned} H(x,y,z) &= \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix} \\ &= \begin{bmatrix} 2x & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}. \end{aligned}$$

- En $(1,1,1)$:

$$H(1,1,1) = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Luego,

$$\begin{aligned} |H(1,1,1) - \lambda I_3| &= \begin{vmatrix} 2-\lambda & -1 & 0 \\ -1 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} \\ &= (2-\lambda)(\lambda^2 - 3\lambda + 1); \end{aligned}$$

entonces los valores propios son $\lambda_1 = 2 > 0$, $\lambda_2 = \frac{3}{2} + \frac{1}{2}\sqrt{5} > 0$ y $\lambda_3 = \frac{3}{2} - \frac{1}{2}\sqrt{5} > 0$. Por tanto, f alcanza un mínimo en $(1,1,1)$.

- En $(0,0,1)$

$$H(0,0,1) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

y

$$\begin{aligned} |H(0,0,1) - \lambda I_3| &= \begin{vmatrix} -\lambda & -1 & 0 \\ -1 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} \\ &= (2-\lambda)(\lambda^2 - \lambda - 1); \end{aligned}$$

así que los valores propios son $\lambda_1 = 2 > 0$, $\lambda_1 = \frac{1}{2} + \frac{1}{2}\sqrt{5} > 0$ y $\lambda_3 = \frac{1}{2} - \frac{1}{2}\sqrt{5} < 0$. Por tanto, f tiene un punto silla en $(0,0,1)$. ✔

6.8.6 Condiciones necesarias para que cierto tipo de funcionales en espacios de dimensión infinita alcancen valores extremos

Vimos en la sección de problemas físicos, que los funcionales que modelan los fenómenos descritos son operadores integrales. Por tanto, al calcular la diferencial de uno de éstos aparece nuevamente el operador integración y, entonces, los puntos críticos deben cumplir ciertas condiciones que tienen que ver con el operador integración y los elementos del espacio donde está contenido el dominio del funcional.

► **Ejemplo 6.77** Recordemos que la curva de longitud mínima que une a dos puntos en el plano es aquella donde el funcional

$$J(f) = \int_a^b \sqrt{1+[f']^2}, \;\; f \in C^1[a,b]$$

alcanza un mínimo. Por el ejemplo 6.65

$$J'(f)h = \int_a^b \frac{f'h'}{\sqrt{1+[f']^2}}$$

y, por tanto, si el problema tiene solución f, entonces debe satisfacer

$$\int_a^b \frac{f'h'}{\sqrt{1+[f']^2}} = 0 \quad \forall h \in C^1[a,b] \blacktriangleleft \tag{6.70}$$

Así, necesitamos encontrar condiciones que nos ayuden a determinar qué funciones cumplen, por ejemplo, con (6.70). Los tres siguientes lemas están diseñados para tal objetivo.

Lemas básicos

Lema 6.1 **(Lagrange)** *Sea $f \in C[a,b]$ una función tal que*

$$\int_a^b fh = 0$$

para toda $h \in C[a,b]$ tal que $h(a) = h(b) = 0$. Entonces $f \equiv 0$ en $[a,b]$.

DEMOSTRACIÓN ■ Supongamos que $f(x_0) \neq 0$ para algún $x_0 \in [a,b]$. Por continuidad de f existe un intervalo $[x_1,x_2] \subset [a,b]$ alrededor de x_0 tal que $f(x)$ tiene el mismo signo que $f(x_0)$ en $[x_1,x_2]$. Sea

$$h(x) = \begin{cases} (x-x_1)(x_2-x) & \text{si } x_1 \le x \le x_2 \\ 0 & \text{en otro caso} \end{cases}$$

Entonces $h \in C[a,b]$, $h(a) = h(b) = 0$ y $f(x)h(x)$ tiene el mismo signo que $f(x)$ en $[x_1,x_2]$; pero $\int_a^b fh = \int_{x_1}^{x_2} fh \neq 0$, lo cual es una contradicción. ■

Lema 6.2 *Sea $f \in C[a,b]$ tal que*

$$\int_a^b fh' = 0 \tag{6.71}$$

para toda $h \in C^1[a,b]$ que satisface $h(a) = h(b) = 0$, entonces f es constante en $[a,b]$.

DEMOSTRACIÓN ■ Note que si se añade la hipótesis $f \in C^1[a,b]$, la demostración es inmediata al integrar por partes y utilizar el lema de Lagrange (lema 6.1). Demostremos este lema sólo suponiendo que f es continua en $[a,b]$. Sea

$$c = \frac{1}{b-a}\int_a^b f.$$

Entonces por (6.71), se tiene

$$\int_a^b (f-c)\,h' = 0 \tag{6.72}$$

para toda $h \in C^1[a,b]$ con $h(a) = h(b) = 0$. Sea

$$h(t) = \int_a^t (f-c)$$

entonces $h'(t) = f(t) - c$ y, por ende, $h' \in C[a,b]$. Además,

$$\begin{aligned} &h(a) = 0 \text{ y} \\ &h(b) = \left(\int_a^b f\right) - c(b-a) = 0 \end{aligned}$$

Entonces, al aplicar a esta h particular (6.72) se obtiene

$$\int_a^b (f-c)(f-c) = 0.$$

Por tanto, el producto interior de $f-c$ consigo misma en el espacio $C[a,b]$ es cero; luego $f = c$. ■

Lema 6.3 **(Du Bois-Reymond)** *Sean $\varphi, \psi \in C[a,b]$ tales que*

$$\int_a^b (\varphi h + \psi h') = 0 \tag{6.73}$$

para toda $h \in C^1[a,b]$ que satisface la condición $h(a) = h(b) = 0$. Entonces $\psi \in C^1[a,b]$ y además $\psi' = \varphi$ en $[a,b]$.

DEMOSTRACIÓN ■ Sea

$$F(x) = \int_a^x \varphi$$

Entonces $F'(x) = \varphi(x)\ \forall x \in [a,b]$; luego $F \in C^1[a,b]$. Por otra parte, para toda $h \in C^1[a,b]$, con $h(a) = h(b) = 0$, se tiene

$$\begin{aligned} \int_a^b \varphi h &= \int_a^b F'h \\ &= Fh|_a^b - \int_a^b Fh' \\ &= -\int_a^b Fh' \end{aligned}$$

y por (6.73) se deduce que

$$\int_a^b (F-\psi)\, h' = 0\,.$$

Del lema 6.2

$$\psi - F = c$$

lo cual prueba que $\psi \in C^1[a,b]$. Finalmente,

$$\psi' = F' = \varphi\,. \quad ■$$

▶ **Ejemplo 6.78** **(Geodésicas en el plano)** Sea $J(f) = \int_a^b \sqrt{1+[f']^2}$, $f \in C^1[a,b]$. Se quiere encontrar los puntos críticos de este funcional que satisfagan la condición $f(a) = A$ y $f(b) = B$. Por el (6.70) del ejemplo 6.77 los puntos críticos de este funcional satisfacen

$$J'(f)\,h = \int_a^b \frac{f'h'}{\sqrt{1+[f']^2}} = 0$$

para toda $h \in C^1[a,b]$. Por el lema 6.2

$$\frac{f'}{\sqrt{1+[f']^2}} = c_1$$

donde c_1 es una constante; de aquí que

$$f' = \alpha$$

donde α es una constante. Por tanto,

$$f(x) = \alpha x + \beta$$

que, como es natural, es la línea recta que pasa por los puntos (a,A) y (b,B), con $\alpha = (B-A)/(b-a)$ y $\beta = (bA - Ba)/(b-a)$.◀

▶ **Ejemplo 6.79** **(Braquistócrona)** Sea $J(f) = \frac{1}{\sqrt{2g}} \int_a^b \sqrt{\frac{1+[f']^2}{f}}$. Se quieren encontrar los puntos críticos de este funcional con la condición $f(a) = A$ y $f(b) = B$. Por el ejemplo 6.67 (ya que el operador integración es lineal y continuo), se debe tener

$$J'(f)\,h = \int_a^b \left[\frac{f'h'}{f^{1/2}\left(1+(f')^2\right)^{1/2}} - \frac{\left(1+(f')^2\right)^{1/2} h}{2f^{3/2}} \right] = 0$$

para todo $h \in C^1[a,b]$. Por el lema de Dubois-Reymond $\dfrac{f'}{f^{1/2}\left(1+(f')^2\right)^{1/2}} \in C^1[a,b]$ y

$$\left(\frac{f'}{f^{1/2}\left(1+(f')^2\right)^{1/2}}\right)' = -\frac{\left(1+(f')^2\right)^{1/2}}{2f^{3/2}} \tag{6.74}$$

Por tanto,

$$\frac{1}{f\left(1+(f')^2\right)}\left[\begin{array}{c} f''f^{1/2}\left(1+(f')^2\right)^{1/2} \\ -f'\left(\begin{array}{c} \frac{1}{2}f^{-1/2}f'\left(1+(f')^2\right)^{1/2} \\ +f^{1/2}\frac{1}{2}\left(1+(f')^2\right)^{-1/2}2f'f'' \end{array}\right)\end{array}\right] = -\frac{\left(1+(f')^2\right)^{1/2}}{2f^{3/2}}$$

luego,

$$\frac{1}{2f^{3/2}\left(1+(f')^2\right)^{3/2}}\left[2f''f\left(1+(f')^2\right)-(f')^2\left(1+(f')^2+2ff''\right)\right] = -\frac{\left(1+(f')^2\right)^{1/2}}{2f^{3/2}}$$

que equivale a

$$2f''f\left(1+(f')^2\right)-(f')^2\left(1+(f')^2+2ff''\right) = -\left(1+(f')^2\right)^2$$

y que se reduce a

$$2f''f+(f')^2+1=0.$$

Si se multiplica la última igualdad en ambos lados por f', se obtiene

$$2f''ff'+\left(1+(f')^2\right)f'=0$$

que equivale a

$$\left[f\left(1+(f')^2\right)\right]'=0$$

y, por tanto,

$$f\left(1+(f')^2\right)=c$$

donde c es una constante, la cual es la ecuación diferencial de la braquistócrona (6.61) que Bernoulli propuso para resolver este problema y que vimos en el apartado de problemas físicos al inicio de esta sección (cfr. pág. 675).◀

Los funcionales que propusimos para resolver los problemas físicos al inicio de esta sección son de una de las formas generales:

1. $J : C[a,b] \to \mathbb{R}$, definido por

$$J(f)=\int_a^b F(x,f(x))\,dx$$

donde F es una función clase C^1 en $[a,b]\times\mathbb{R}$; esto es, una función diferenciable con continuidad en $[a,b]\times\mathbb{R}$.

2. $J : C^1[a,b] \to \mathbb{R}$, definido como

$$J(f) = \int_a^b F(x, f(x), f'(x))\, dx \tag{6.75}$$

donde F es una función clase C^1 en $[a,b]\times\mathbb{R}$; esto es, una función diferenciable con continuidad en $[a,b]\times\mathbb{R}$.

Es posible establecer condiciones necesarias generales para que este par de funcionales alcancen valores óptimos.

Condiciones necesarias para que el funcional $J(f) = \int_a^b F(x, f(x))\, dx$ alcance un extremo relativo

Supongamos que $F : [a,b]\times\mathbb{R} \to \mathbb{R}$, es una función diferenciable con continuidad en $[a,b]\times\mathbb{R}$. Se define $J : C[a,b] \to \mathbb{R}$ como

$$J(f) = \int_a^b F(x, f(x))\, dx \tag{6.76}$$

Sea $J_1 : C[a,b] \to \mathbb{R}$, $f \mapsto J_1(f)$, donde

$$J_1(f)(x) = F(x, f(x)).$$

Si $f, h \in C[a,b]$ y $x \in [a,b]$ son fijos, entonces

$$\begin{aligned}\lim_{t\to 0} \frac{F(x, f(x)+th(x)) - F(x, f(x))}{t} &= F'(x, f(x))h(x)\\ &= \frac{\partial F}{\partial x_2}(x, f(x))h(x).\end{aligned}$$

Sea $L : C[a,b] \to C[a,b]$ la transformación definida, para cada $y \in [a,b]$, como

$$(Lu)y = \frac{\partial F}{\partial x_2}(y, f(y))u(y).$$

$Lu \in C[a,b]$ porque $\dfrac{\partial F}{\partial x_2}$, f y u son funciones continuas en $[a,b]$; claramente L es lineal. Como la función $y \mapsto \dfrac{\partial F}{\partial x_2}(y, f(y))$ es continua en $[a,b]$, entonces es acotada y, por tanto, existe $M > 0$ tal que

$$\left|\frac{\partial F}{\partial x_2}(y, f(y))\right| \le M$$

para todo $y \in [a,b]$. Así que, para todo $y \in [a,b]$,

$$\begin{aligned}|(Lu)y| &= \left|\frac{\partial F}{\partial x_2}(y, f(y))u(y)\right|\\ &\le M\,|u(y)|\\ &\le M\,\|u\|_\infty\end{aligned}$$

y, por ende,

$$\|Lu\|_\infty \leq M\,\|u\|_\infty$$

por lo que L es continua. Por el teorema del valor medio existe c entre $f(x)$ y $f(x)+h(x)$ tal que

$$F(x, f(x)+h(x)) - F(x, f(x)) = \frac{\partial F}{\partial x_2}(x,c)h(x).$$

Por tanto,

$$\begin{aligned}
\left| F(x, f(x)+h(x)) - F(x, f(x)) - \frac{\partial F}{\partial x_2}(x, f(x))h(x) \right| &= \left| \frac{\partial F}{\partial x_2}(x, f(x))h(x) - \frac{\partial F}{\partial x_2}(x,c)h(x) \right| \\
&= \left| \frac{\partial F}{\partial x_2}(x, f(x)) - \frac{\partial F}{\partial x_2}(x,c) \right| |h(x)| \\
&\leq \left| \frac{\partial F}{\partial x_2}(x, f(x)) - \frac{\partial F}{\partial x_2}(x,c) \right| \|h\|_\infty .
\end{aligned}$$

Lo cual implica

$$\frac{\left| F(x, f(x)+h(x)) - F(x, f(x)) - \frac{\partial F}{\partial x_2}(x, f(x))h(x) \right|}{\|h\|_\infty} \leq \left| \frac{\partial F}{\partial x_2}(x, f(x)) - \frac{\partial F}{\partial x_2}(x,c) \right|.$$

Ya que $\dfrac{\partial F}{\partial x_2}$ es continua, dado $\varepsilon > 0$ existe δ tal que

$$\|(x,y),(x,y_1)\| < \delta \Rightarrow \left| \frac{\partial F}{\partial x_2}(x,y) - \frac{\partial F}{\partial x_2}(x,y_1) \right| < \varepsilon.$$

Entonces,

$$\begin{aligned}
\|h\|_\infty < \delta \quad &\Rightarrow \quad |f(x) - c| = \|(x, f(x)) - (x,c)\|_\infty < \delta \\
&\Rightarrow \quad \left| \frac{\partial F}{\partial x_2}(x, f(x)) - \frac{\partial F}{\partial x_2}(x,c) \right| < \varepsilon.
\end{aligned}$$

Esto es,

$$\|h\|_\infty < \delta \Rightarrow \frac{\|J_1(f+h) - J_1(f) - Lh\|_\infty}{\|h\|_\infty} < \varepsilon$$

es decir,

$$\lim_{h\to 0} \frac{J_1(f+h) - J_1(f) - Lh}{\|h\|_\infty} = 0.$$

Por tanto, J_1 es diferenciable en f y $J_1'(f) = L$. Por la regla de la cadena y el ejemplo 6.61 J es diferenciable en todo $f \in C[a,b]$ y

$$J'(f)h = \int_a^b \frac{\partial F}{\partial x_2}(x, f(x))h(x)dx.$$

Entonces, si J alcanza un extremo relativo en f se debe tener

$$\int_a^b \frac{\partial F}{\partial x_2}(x, f(x))h(x)dx = 0$$

para todo $h \in C[a,b]$. Esto implica la función $x \mapsto \frac{\partial F}{\partial x_2}(x, f(x))$ es ortogonal a toda función $h \in C[a,b]$ y, por ende,

$$\frac{\partial F}{\partial x_2}(x, f) = 0\,.$$

Esta es la condición necesaria para que el funcional (6.76) alcance un extremo en f.

Condición de Euler

Ahora investiguemos condiciones necesarias para que el funcional

$$J(f) = \int_a^b F(x, f(x), f'(x))\,dx, \tag{6.77}$$

donde $f \in C^1[a,b]$ y F es una función clase C^1, alcance un extremo. De manera similar al caso anterior, podemos mostrar que la diferencial de la función $f \mapsto J_1(f)$, donde $J_1(f)(x) = F(x, f(x), f'(x))$ es diferenciable y que la diferencial de J_1 es la aplicación $h \mapsto J_1(f)h$, con

$$J_1'(f)\,h(x) = F_y(x, f(x), f'(x))\,h(x) + F_z(x, f(x), f'(x))\,h'(x)\,.$$

Por lo que J es diferenciable y, por la regla de la cadena y el ejemplo 6.61,

$$J'(f)\,h = \int_a^b [F_y(x, f, f')\,h + F_z(x, y, y')\,h']\,dx\,.$$

Por ello, si J alcanza un extremo en f, entonces debe cumplir

$$\int_a^b [F_y(x, f, f')\,h + F_z(x, y, y')\,h']\,dx = 0\ \forall\ h \in C^1[a,b]\,.$$

Por el lema de Dubois-Reymond se debe tener

$$\frac{\partial}{\partial f}F(x, f, f') - \frac{d}{dx}\left[\frac{\partial}{\partial f'}F(x, f, f')\right] = 0 \tag{6.78}$$

La relación (6.78) es conocida como **condición necesaria de Euler** para que el funcional (6.77) alcance un extremo relativo.

▶ **Ejemplo 6.80** **(Distancia mínima)** En el problema de la distancia mínima entre dos puntos en el plano, el funcional a minimizar es

$$J'(f) = \int_a^b \sqrt{1 + (f')^2}\,.$$

Entonces, si $f \in C^1[a,b]$ es un mínimo de J, de la condición (6.78), se debe cumplir

$$\frac{\partial}{\partial f}\left(\sqrt{1 + (f')^2}\right) - \frac{d}{dx}\left(\frac{\partial}{\partial f'}\left(\sqrt{1 + (f')^2}\right)\right) = 0;$$

esto es,

$$\frac{d}{dx}\left(\frac{\partial}{\partial f'}\left(\sqrt{1+(f')^2}\right)\right)=0$$

y, por tanto,

$$\frac{f'}{\sqrt{1+(f')^2}}=c_1$$

donde c_1 es una constante; luego

$$f(x)=c_1x+c_2\,.$$◄

► **Ejemplo 6.81** **(Braquistócrona)** Nuevamente si consideramos el problema de la braquistócrona, el funcional a minimizar es

$$J(f)=\frac{1}{\sqrt{2g}}\int_a^b\sqrt{\frac{1+(f')^2}{f}}$$

La condición de Euler (6.78) para este caso es

$$\frac{\partial}{\partial f}\left(\sqrt{\frac{1+(f')^2}{f}}\right)-\frac{d}{dx}\left(\frac{\partial}{\partial f'}\left(\sqrt{\frac{1+(f')^2}{f}}\right)\right)=0$$

$$\therefore\quad -\frac{1}{2f^{3/2}}\sqrt{1+(f')^2}-\frac{d}{dx}\left(\frac{1}{2}\frac{1}{\sqrt{f}}\left(1+(f')^2\right)^{-1/2}2f'\right)=0$$

esto es

$$\left(\frac{f'}{f^{1/2}\left(1+(f')^2\right)^{1/2}}\right)'=-\frac{\left(1+(f')^2\right)^{1/2}}{2f^{3/2}}$$

que es la relación (6.74) del ejemplo 6.79, donde vimos que conduce a la ecuación diferencial de la braquistócrona (6.61)

$$f\left(1+(f')^2\right)=c\,.$$◄

► **Ejemplo 6.82** Hallar las funciones donde el funcional

$$J(f)=\int_0^1(16f^2(x)+16f(x)f'(x)+4[f'(x)]^2)dx,$$

con la condición de frontera

$$\begin{aligned} f(0)&=0\\ f(1)&=e^2-e^{-2}\end{aligned}\qquad(6.79)$$

alcanza valores extremos.◄

Solución En este caso $F(x,f,f') = 16f^2 + 16ff' + 4(f')^2$. Entonces,

$$\frac{\partial F}{\partial f} = 32f + 16f',$$
$$\frac{\partial F}{\partial f'} = 16f + 8f' \text{ y}$$
$$\frac{d}{dx}\left(\frac{\partial F}{\partial f'}\right) = 16f' + 8f''.$$

Por tanto, la ecuación de Euler que corresponde es

$$32f + 16f' = 16f' + 8f''$$

que equivale a

$$f'' - 4f = 0 \tag{6.80}$$

Utilizando el ejercicio propuesto 191 de este capítulo (cfr. pág. 751), el polinomio característico es

$$\lambda^2 - 4 = 0$$

que tiene por raíces a $\lambda_1 = 2$ y $\lambda_2 = -2$, con lo que la solución general de (6.80) está dada por

$$f = C_1 e^{2x} + C_2 e^{-2x}.$$

Como $f(0) = 0$ y $f(1) = e^2 - e^{-2}$, entonces se debe cumplir

$$C_1 + C_2 = 0$$
$$C_1 e^2 + C_2 e^{-2} = e^2 - e^{-2}$$

y al resolver el sistema obtenemos $C_1 = 1$ y $C_2 = -1$. Por tanto, J alcanza un extremo en

$$f_0(x) = e^{2x} - e^{-2x} = 2\operatorname{senh}(2x),\ x \in [0,1] \tag{6.81}$$

el cual puede ser máximo o mínimo o punto de silla. Sin embargo, dado que la solución de (6.80) con las condiciones de frontera (6.79) es única, y

$$J(f) = \int_0^1 (4f + 2f')^2 \geq 0,$$

se desprende que en (6.81) J alcanza máximo relativo y, de hecho, absoluto. El valor mínimo, si existe, se alcanza en la solución de

$$4f + 2f' = 0$$

con la condición de frontera (6.79). Pero la anterior ecuación equivale a

$$\frac{df}{dx} = -2f$$

que al separar variables produce

$$\frac{df}{f} = -2$$

y al integrar y reducir se obtiene

$$f(x) = C_1 e^{-2x} \text{ o } f = 0.$$

Ninguna de estas funciones, para cualquier valor de C_1, cumple con la condición de frontera $f(0) = 0$, $f(1) = e^2 - e^{-2}$. Por ende, J no tiene mínimos relativos ni absolutos en $C^1[0,1]$ que satisfagan la condición de frontera (6.79). ✔

6.8.7 Dinámica de un monopolista

Supongamos[24] que una empresa monopolista produce una sola mercancía, con un costo total por unidad de tiempo que es una función cuadrática de la cantidad producida:

$$q(u) = Au^2 + Bu + C$$

donde A, B y C son constantes positivas. Se supone que la cantidad demandada depende del precio $p(t)$ y de su razón instantánea de cambio $p'(t)$ en la forma

$$y = ap + b + hp',$$

donde $a < 0$, $b > 0$ y $h \neq 0$, y que toda la mercancía que se produce es consumida por el mercado. Entonces, la ganancia de la empresa es

$$\begin{aligned}
\pi &= pu - q \\
&= py - q \\
&= p(ap + b + hp') - (Au^2 + Bu + C) \\
&= p(ap + b + hp') - A(ap + b + hp')^2 - B(ap + b + hp') - C.
\end{aligned}$$

El objetivo es hacer máxima la ganacia en un intervalo de tiempo $[0, T]$; por tanto, se debe hallar el valor máximo del funcional

$$J(p) = \int_0^T \big(p(ap + b + hp') - A(ap + b + hp')^2 - B(ap + b + hp') - C\big)\, dt \tag{6.82}$$

sujeto a la condición de frontera

$$\begin{aligned}
p(0) &= p_1 \\
p(T) &= p_2
\end{aligned} \tag{6.83}$$

En este caso

$$F(t, p, p') = p(ap + b + hp') - A(ap + b + hp')^2 - B(ap + b + hp') - C$$

entonces,

$$\frac{\partial F}{\partial p} = 2ap + hp' - 2Aa(ap + b + hp') - Ba,$$

[24]Cfr. *The American Mathematical Monthly*, vol. 31, núm. 2 (febrero, 1924), pp. 77-83.

$$\begin{aligned}\frac{\partial F}{\partial p'} &= hp - 2Ah(ap + b + hp') - Bh \qquad \text{y}\\ \frac{d}{dx}\left(\frac{\partial F}{\partial p}\right) &= hp' - 2Ah(ap' + hp'').\end{aligned}$$

La condición de Euler para este problema es

$$hp' - 2Ah(ap' + hp'') = 2ap + hp' - 2Aa(ap + b + hp') - Ba$$

que se reduce a

$$2Ah^2p'' + 2a(1 - Aa)p = Ba + 2Aab - b \tag{6.84}$$

El polinomio característico (cfr. problema propuesto 198) está dado por

$$2Ah^2\lambda^2 + 2a(1 - Aa) = 0$$

cuyas raíces son[25]

$$\lambda = \pm\frac{1}{|h|}\sqrt{\frac{a(Aa-1)}{A}}\,.$$

Así, la solución de la ecuación diferencial homogénea asociada a la ecuación (6.84) es

$$p_H(t) = C_1e^{\lambda t} + C_2e^{-\lambda t}.$$

Si $p = p_0$ es una solución constante de (6.84), entonces

$$2a(1 - Aa)p_0 = Ba + 2Aab - b$$

y, por tanto,[26]

$$p_0 = \frac{Ba + 2Aab - b}{2a(1 - Aa)} \tag{6.85}$$

por lo que la solución de (6.84) es

$$p^*(t) = C_1e^{\lambda t} + C_2e^{-\lambda t} + p_0 \tag{6.86}$$

La solución (6.86) debe, además, satisfacer las condiciones de frontera (6.83). Se obtiene entonces el sistema lineal

$$\begin{aligned}C_1 + C_2 &= p_1 - p_0\\ C_1e^{\lambda T} + C_2e^{-\lambda T} &= p_2 - p_0\end{aligned}$$

que tiene solución única

$$C_1 = \frac{p_1 - p_0 - (p_2 - p_0)e^{\lambda T}}{1 - e^{2\lambda T}}$$

[25] Note que $a(Aa - 1) = Aa^2 - a > 0$ porque $a < 0$ y $A > 0$.

[26] Observe que p_0 es positivo porque $a < 0$ y las demás constantes son positivas.

y

$$C_2 = \frac{p_1 - p_0 - (p_2 - p_0)e^{-\lambda T}}{1 - e^{-2\lambda T}}.$$

Hemos probado que si el funcional (6.82) sujeto a las condiciones de frontera (6.83) tiene un máximo, entonces éste debe alcanzar p^*, donde $p^*(t)$ está dado por (6.86) y, C_1, C_2 son los valores precedentes. Desafortunadamente, no podemos determinar, con la información que tenemos, si en p^* se alcanza un máximo, o si en realidad ahí se tiene un mínimo o un punto de silla. Para ello es necesario conocer más información acerca de los parámetros del problema.

6.8.8 Epílogo

Dar condiciones suficientes para determinar si en un punto crítico se alcanza un máximo o un mínimo o un punto silla, de un funcional, en un espacio normado de dimensión infinita, que se apliquen con relativa facilidad —como el criterio de signos de los valores propios de la matriz hessiana para el caso de dimensión finita— al menos para los problemas que hemos visto en esta sección, es complejo y no lo trataremos en este libro. En el problema planteado por Evans y en el de la braquistócrona, únicamente encontramos que en ambos existe un único punto que es el candidato, para que en él se alcance un valor óptimo. Pero no demostramos que en el primero se tiene un máximo y en el segundo un mínimo. En el caso de la braquistócrona Newton y J. Bernoulli tampoco lo hicieron; esto no es de ninguna manera un consuelo. Lo único que se quiere mostrar es que la naturaleza de este tipo de problemas es muy delicada y que, en general, en las aplicaciones surgen dificultades como éstas. Más aún, seguramente el lector puede hacer una crítica objetiva de la manera de resolver estos problemas; por ejemplo el de la braquistócrona. En él se puede cuestionar si efectivamente existe conservación de energía; o si el símil del deslizamiento del abalorio con el movimiento de la luz, a través de medios con distinta densidad, dado por Bernoulli es correcto; o por qué se evitó la mención de la discontinuidad en cero del integrando del funcional que se desea minimizar y, por tanto, la probable no convergencia de la integral; o por qué se derivó el integrando considerando $f \neq 0$ en todo el intervalo $[0,b]$. Las dos primeras cuestiones tienen que ver con el planteamiento físico y con la naturaleza y filosofía de esta ciencia (la Física) y sólo la contrastación experimental puede validar si Newton y Bernoulli tenían razón.[27] Las últimas dos son dificultades técnicas de carácter matemático que se pueden soslayar cambiando el conjunto de funciones admisibles; es decir, el espacio vectorial y la norma; algo que no se hizo para evitar distraer al lector con más detalles y tratar el problema en el marco matemático más simple que fue posible. Sin embargo, sí existen criterios matemáticos que dan condiciones suficientes para determinar la naturaleza de los puntos críticos para los funcionales que hemos estudiado en esta sección y para otros que se presentan en muchas aplicaciones. Se invita al lector a consultar la bibliografía dada en este libro para conocerlos y, de ser posible, estudiarlos. Aquí, mediante una transición natural del concepto de derivación para funciones de una variable real al concepto de diferencial de funciones en espacios vectoriales, mostrando cómo adquieren relevancia los términos de espacio vectorial, normas, transformaciones lineales, etc., se buscó introducir en forma muy elemental al lector en el fascinante tema del cálculo de variaciones como una excelente razón para ilustrar la importancia y relación que tiene el análisis lineal (álgebra lineal) con otras ramas avanzadas de las mátematicas y sus aplicaciones.

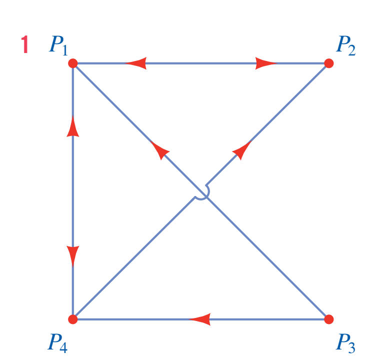
[27] Cfr. *An experimental study of the brachistochrone*, F. M. Phelps III, *et al.*, 1982, Eur. J. Phys. **3** 1-4.

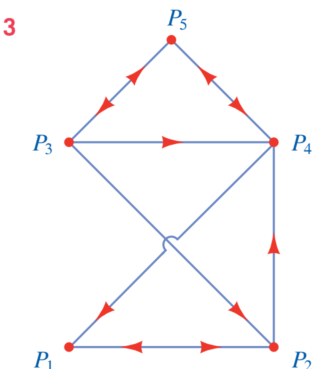
6.9 Ejercicios propuestos

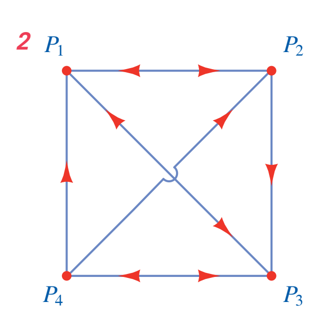
El lector encontrará la respuesta a los ejercicios en cursiva en el apéndice E al final del libro.

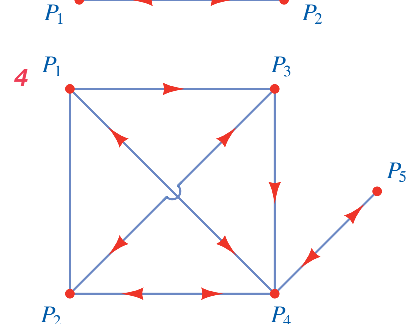
Matrices de incidencia y teoría de grafos (respuestas en páginas 1092-1093)

En los ejercicios 1 a 4 encontrar la matriz de incidencia A para el digrafo dado.

1


3


2


4


5 Utilizar el teorema 6.1 para encontrar el número de formas distintas en las que P_3 puede enviar un mensaje, con dos intermediarios, a P_2 para el digrafo del ejercicio 1 y escribir las 3-cadenas correspondientes.

6 Utilizar el teorema 6.1 para encontrar el número de formas distintas en las que P_3 puede enviar un mensaje, con dos intermediarios, a P_1 para el digrafo del ejercicio 1 y escribir las 3-cadenas correspondientes.

7 Utilizar el teorema 6.1 para encontrar el número de formas distintas en las que P_1 puede enviar un mensaje, con dos intermediarios, a P_2 para el digrafo del ejercicio 2 y escribir las 3-cadenas correspondientes.

8 Utilizar el teorema 6.1 para encontrar el número de formas distintas en las que P_1 puede enviar un mensaje, con dos intermediarios, a P_4 para el digrafo del ejercicio 2 y escribir las 3-cadenas correspondientes.

9 Utilizar el teorema 6.1 para encontrar el número de formas distintas en las que P_3 puede enviar un mensaje, con un intermediario, a P_4 para el digrafo del ejercicio 3 y escribir las 2-cadenas correspondientes.

10 Utilizar el teorema 6.1 para encontrar el número de formas distintas en las que P_5 puede enviar un mensaje, con un intermediario, a P_2 para el digrafo del ejercicio 3 y escribir las 2-cadenas correspondientes.

11 Encontrar el número de formas distintas en las que P_2 puede transmitir a P_3 un mensaje en el digrafo del ejercicio 4 en a lo más 3 etapas (cuando mucho con dos intermediarios).

12 Encontrar el número de formas distintas en las que P_4 puede transmitir a P_3 un mensaje en el digrafo del ejercicio 4 en a lo más 3 etapas (cuando mucho con dos intermediarios).

13 Encontrar el número de formas distintas en las que P_4 puede transmitir a P_3 un mensaje en el digrafo del ejercicio 4 en a lo sumo 2 etapas (a lo más un intermediario).

14 Mostrar, utilizando el teorema 6.2 (cfr. pág. 585), que el digrafo del ejercicio 4 no tiene clanes.

15 Mostrar, utilizando el teorema 6.2 (cfr. pág. 585), que el digrafo del ejercicio 2 no tiene clanes.

16 Sea la matriz de incidencia

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

(*a*) Encontrar la matriz asociada $B = \mathscr{A}(A)$.
(*b*) Bosquejar el digrafo de la matriz A.
(*c*) Utilizar el teorema 6.2 para probar que todo punto del digrafo pertenece a un clan.
(*d*) Hallar los clanes del digrafo.

17 Sea la matriz de incidencia

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

(*a*) Encontrar la matriz asociada $B = \mathscr{A}(A)$.
(*b*) Bosquejar el digrafo de la matriz A.
(*c*) Utilizar el teorema 6.2 para probar que todo punto del digrafo pertenece a un clan.
(*d*) Hallar los clanes del digrafo.

18 Sea la matriz de incidencia

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

(*a*) Encontrar la matriz asociada $B = \mathscr{A}(A)$.
(*b*) Bosquejar el digrafo de la matriz A.
(*c*) Utilizar el teorema 6.2 para probar que todo punto del digrafo pertenece a un clan.
(*d*) Hallar los clanes del digrafo.

19 Sea la matriz de incidencia

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

(a) Probar que A corresponde a una relación de dominancia total.

(b) Determinar las relaciones de dominancia entre los individuos P_i en a lo más dos etapas.

20 Sea la matriz de incidencia

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

(a) Probar que A corresponde a una relación de dominancia total.

(b) Determinar las relaciones de dominancia entre los individuos P_i en a lo sumo dos etapas.

21 Sea la matriz de incidencia

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

(a) Probar que A corresponde a una relación de dominancia total.

(b) Determinar las relaciones de dominancia entre los individuos P_i en dos etapas cuando mucho.

22 Encontrar el menor número de etapas que necesita P_1 para dominar a P_2 en el ejercicio anterior.

23 Sea la matriz de incidencia

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

(a) Probar que A corresponde a una relación de dominancia parcial.

(b) Determinar las relaciones de dominancia entre los individuos P_i en dos etapas a lo sumo.

24 Sea la matriz de incidencia

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

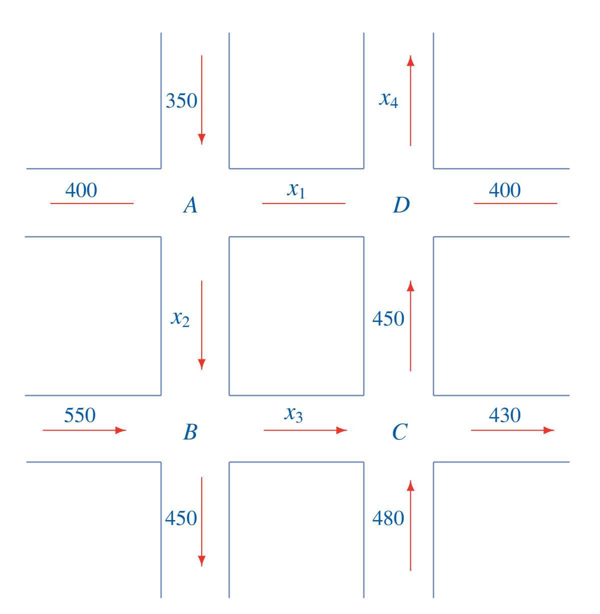
(a) Probar que A corresponde a una relación de dominancia parcial.

(b) Determinar las relaciones de dominancia entre los individuos P_i en a lo más dos etapas.

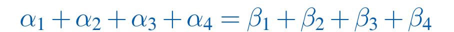
Redes de conducción y principios de conservación (respuestas en páginas 1093-1094)

Flujo vehicular y redes de conducción

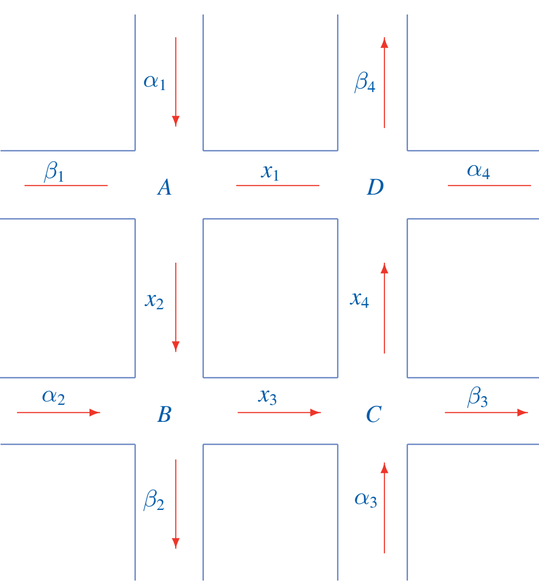
25 Determinar los valores x_j, $j = 1, 2, 3, 4$, del siguiente diagrama de flujo vehicular:



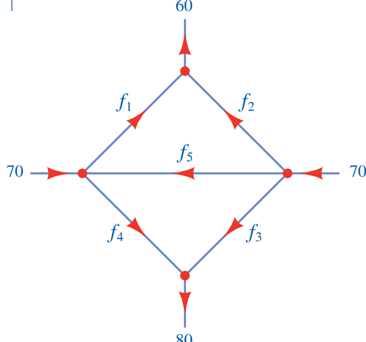
26 Demostrar que el sistema lineal que corresponde al problema de encontrar los valores x_j, $j = 1, 2, 3, 4$, del diagrama de flujo vehicular contenido en la siguiente figura es consistente si y sólo si la cantidad de vehículos que entra al circuito es igual a la cantidad de vehículos que sale del mismo; esto es,



$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = \beta_1 + \beta_2 + \beta_3 + \beta_4$$

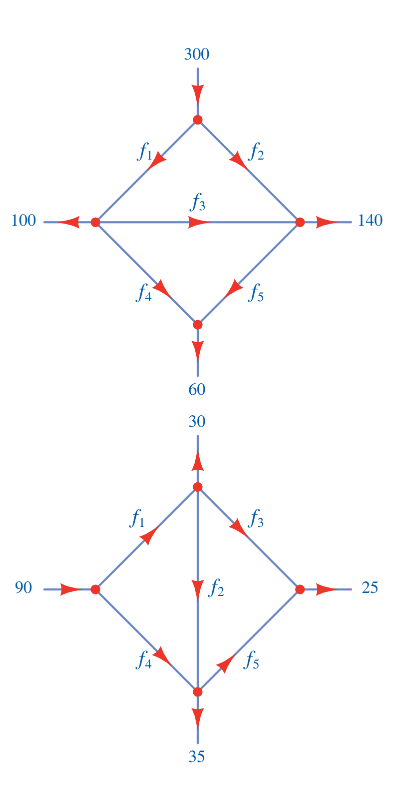


27 *(a)* Hallar todas las posibles magnitudes de los flujos f_j de la siguiente red conductora.
(b) Si no existe la posibilidad de cambiar las direcciones en cada uno de los conductores de la red, determinar las condiciones que deben cumplir las magnitudes de los flujos f_j.



28 La red del siguiente diagrama conduce agua; las unidades de los flujos son en litros por minuto.

(a) Encontrar todos los posibles valores de los flujos f_j.
(b) Si la tubería que corresponde al flujo f_4 es cerrada, hallar los valores de los flujos f_j de la red.
(c) En las condiciones del inciso anterior, determinar el valor mínimo de f_1 si las direcciones de los flujos en el diagrama no se pueden invertir.



29 La red del siguiente diagrama conduce agua con flujos en unidades de litros por minuto. Por cuestiones técnicas no es posible invertir las direcciones de los flujos que se muestran en ese esquema.

(a) Determinar los posibles valores de los flujos f_i (con las respectivas restricciones).
(b) Si la tubería que corresponde a f_3 se cierra y se debe mantener en cada tubería un flujo no superior a 40 litros por minuto, encontrar los valores que deben tener los flujos f_5 y f_4.

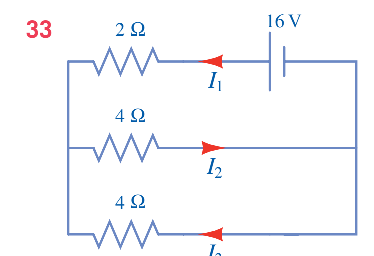
Circuitos eléctricos

En los ejercicios 30 a 37 encontrar las corrientes I_j del circuito dado.

30

10 Ω, 6 V, I_1; 5 Ω, 10 V, I_2; 20 Ω, 10 V, I_3

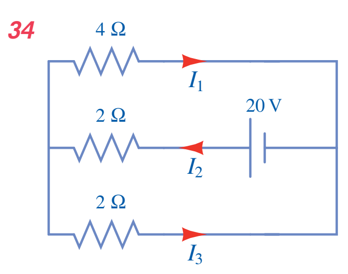
33



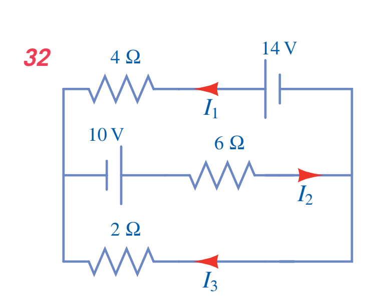
31

8 Ω, 6 V, I_1; 4 Ω, I_2; 8 Ω, 10 V, I_3

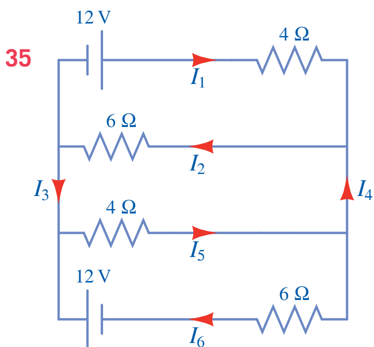
34



32

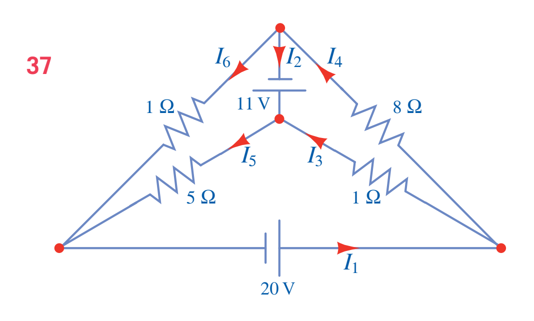


35

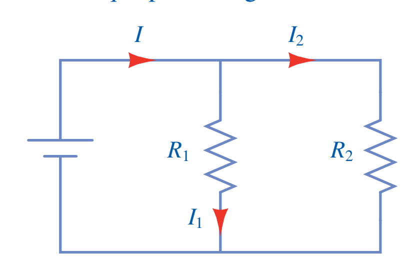


36

37



38 Probar que para el siguiente circuito se tiene:

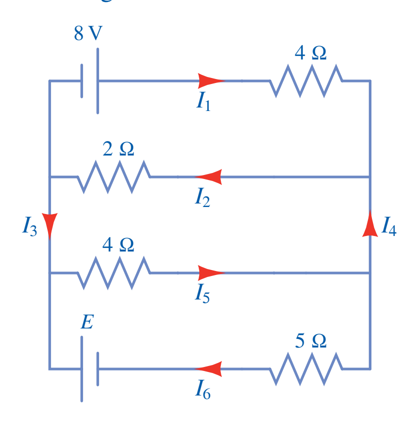


(a) $I_1 = \left(\dfrac{R_2}{R_1+R_2}\right)I = \dfrac{R}{R_1}I$ *(b)* $I_2 = \left(\dfrac{R_1}{R_1+R_2}\right)I = \dfrac{R}{R_2}I$

donde

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}.$$

39 En el siguiente circuito



Hallar la fuerza electromotriz E y las corrientes I_j si se sabe que $I_1 = 2$ A.

Balance químico

En los ejercicios 40 a 49 balancear las ecuaciones químicas dadas.

40 $Na + H_2O \longrightarrow NaOH + H_2$.
41 $Fe + O_2 \longrightarrow Fe_2O_3$.
42 $MnO_2 + KOH + O_2 \longrightarrow K_2MnO_4 + H_2O$.
43 $K_2MnO_4 + CO_2 + H_2O \longrightarrow KMnO_4 + MnO_2 + KHCO_3$.
44 $C_2H_4 + O_2 \longrightarrow CO_2 + H_2O$.
45 $MnO_2 + HCl \longrightarrow MnCl_2 + H_2O + Cl_2$.
46 $C + HNO_3 \longrightarrow NO_2 + CO_2 + H_2O$.
47 $CO + Fe_2O_3 \longrightarrow Fe + CO_2$.
48 $HCl + HClO_3 \longrightarrow H_2O + Cl_2$.
49 $FeCl_3 + H_2S \longrightarrow FeCl_2 + HCl + S$.

Análisis insumo-producto (respuestas en página 1094)

Modelo para economía abierta

50 Una economía hipotética consta de dos industrias, I_1, I_2, con la siguiente tabla de insumo-demanda-producción:

Productor	Consumidor I_1	I_2	Demanda final	Producción total
I_1	12	6	6	24
I_2	6	18	12	36

Encontrar el correspondiente vector de producción $\vec{x}$ si la demanda cambia a:

(a) $\vec{h} = (20, 30)$. *(b)* $\vec{h} = (35, 50)$.

Las cifras están en millones de unidades monetarias.

51 Una economía hipotética consta de dos industrias, I_1, I_2, con la siguiente tabla de insumo-demanda-producción:

Productor	Consumidor I_1	I_2	Demanda final	Producción total
I_1	15	10	5	30
I_2	5	20	15	40

Encontrar el correspondiente vector de producción $\vec{x}$ si la demanda cambia a:

(a) $\vec{h} = (10, 20)$. *(b)* $\vec{h} = (25, 50)$.

Las cifras están en millones de unidades monetarias.

52 Una economía hipotética consta de dos industrias, I_1, I_2, con la siguiente tabla de insumo-demanda-producción:

Productor	Consumidor I_1	I_2	Demanda final	Producción total
I_1	7	12	9	28
I_2	14	6	16	36

Encontrar el correspondiente vector de producción $\vec{x}$ si la demanda cambia a:

(a) $\vec{h} = (33, 44)$. *(b)* $\vec{h} = (44, 22)$.

Las cifras están en millones de unidades monetarias.

53 Una economía hipotética consta de dos industrias, I_1, I_2, con la siguiente tabla de insumo-demanda-producción:

Productor	Consumidor I_1	I_2	Demanda final	Producción total
I_1	3	20	13	36
I_2	18	10	12	40

Encontrar el correspondiente vector de producción $\vec{x}$ si la demanda cambia a:

(a) $\vec{h} = (42, 63)$. *(b)* $\vec{h} = (84, 126)$.

Las cifras están en millones de unidades monetarias.

54 Una economía hipotética consta de tres industrias, I_1, I_2 e I_3, con la siguiente tabla de insumo-demanda-producción:

Productor	Consumidor I_1	I_2	I_3	Demanda final	Producción total
I_1	80	100	200	20	400
I_2	80	100	100	40	320
I_3	80	60	100	60	300

Encontrar el correspondiente vector de producción $\vec{x}$ si la demanda cambia a:

(a) $\vec{h} = (33, 66, 99)$. *(b)* $\vec{h} = (66, 132, 198)$.

Las cifras están en millones de unidades monetarias.

55 Una economía hipotética consta de tres industrias, I_1, I_2 e I_3, con la siguiente tabla de insumo-demanda-producción:

Productor	Consumidor I_1	I_2	I_3	Demanda final	Producción total
I_1	5	20	20	15	60
I_2	15	10	15	40	80
I_3	10	20	10	20	60

Encontrar el correspondiente vector de producción $\vec{x}$ si la demanda cambia a:

(a) $\vec{h} = (46, 92, 46)$. *(b)* $\vec{h} = (92, 138, 184)$.

Las cifras están en millones de unidades monetarias.

56 Una economía hipotética consta de tres industrias, I_1, I_2 e I_3, con la siguiente tabla de insumo-demanda-producción:

Productor	Consumidor I_1	I_2	I_3	Demanda final	Producción total
I_1	12	15	15	6	48
I_2	12	30	9	9	60
I_3	12	15	15	3	45

Encontrar el correspondiente vector de producción $\vec{x}$ si la demanda cambia a:

(a) $\vec{h} = (23, 46, 23)$. *(b)* $\vec{h} = (46, 69, 46)$.

Las cifras están en millones de unidades monetarias.

57 Una economía está dividida en dos sectores: I_1, I_2. La siguiente tabla contiene la información (hipotética) del intercambio de las cantidades de insumos entre ellas por unidad de producción:

Productor	Consumidor I_1	I_2
I_1	1/2	1/4
I_2	1/4	1/2

Encontrar las cantidades que tienen que producir (en millones de unidades monetarias) estas industrias para satisfacer la demanda final:

- 60 unidades del sector I_1.
- 120 unidades del sector I_2.

58 Una economía está dividida en dos sectores: I_1, I_2. La siguiente tabla contiene la información (hipotética) del intercambio de las cantidades de insumos entre ellas por unidad de producción:

Productor	Consumidor I_1 I_2
I_1 I_2	$\begin{bmatrix} 1/2 & 1/6 \\ 1/4 & 11/36 \end{bmatrix}$

Encontrar las cantidades que tienen que producir (en millones de unidades monetarias) estas industrias para satisfacer la demanda final:

- 44 unidades del sector I_1.
- 66 unidades del sector I_2.

59 Una economía está dividida en dos sectores: I_1, I_2. La siguiente tabla contiene la información (hipotética) del intercambio de las cantidades de insumos entre ellas por unidad de producción:

Productor	Consumidor I_1 I_2
I_1 I_2	$\begin{bmatrix} 1/5 & 3/8 \\ 2/5 & 1/4 \end{bmatrix}$

Encontrar las cantidades que tienen que producir (en millones de unidades monetarias) estas industrias para satisfacer la demanda final:

- 72 unidades del sector I_1.
- 90 unidades del sector I_2.

60 Una economía está dividida en tres sectores: I_1, I_2 e I_3. La siguiente tabla contiene la información (hipotética) del intercambio de las cantidades de insumos entre ellas por unidad de producción:

Productor	Consumidor I_1 I_2 I_3
I_1 I_2 I_2	$\begin{bmatrix} 1/4 & 1/3 & 1/4 \\ 1/4 & 1/6 & 1/2 \\ 1/4 & 1/3 & 1/4 \end{bmatrix}$

Encontrar las cantidades que tienen que producir (en millones de unidades monetarias) estas industrias para satisfacer la demanda final:

- 40 unidades del sector I_1.
- 80 unidades del sector I_2.
- 64 unidades del sector I_3.

61 Una economía está dividida en tres sectores: I_1, I_2 e I_3. La siguiente tabla contiene la información (hipotética) del intercambio de las cantidades de insumos entre ellas por unidad de producción:

Productor	Consumidor I_1 I_2 I_3
I_1 I_2 I_2	$\begin{bmatrix} 2/5 & 2/5 & 1/5 \\ 1/5 & 3/10 & 1/10 \\ 1/10 & 1/10 & 3/10 \end{bmatrix}$

Encontrar las cantidades que tienen que producir (en millones de unidades monetarias) estas industrias para satisfacer la demanda final:

- 42 unidades del sector I_1.
- 84 unidades del sector I_2.
- 21 unidades del sector I_3.

62 Una economía está dividida en tres sectores: I_1, I_2 e I_3. La siguiente tabla contiene la información (hipotética) del intercambio de las cantidades de insumos entre ellas por unidad de producción:

Productor	Consumidor I_1	I_2	I_3
I_1	1/5	1/5	2/5
I_2	1/5	2/5	1/5
I_2	2/5	1/5	1/5

Encontrar las cantidades que tienen que producir (en millones de unidades monetarias) estas industrias para satisfacer la demanda final:

- 24 unidades del sector I_1.
- 72 unidades del sector I_2.
- 48 unidades del sector I_3.

Modelo para economía cerrada

63 En una economía sencilla la distribución insumo-producto se lleva de acuerdo con la siguiente tabla (hipotética):

Consumidor	Productor		
	A	B	C
A	0	2/5	2/5
B	3/5	1/5	2/5
C	2/5	2/5	1/5

(a) Encontrar qué condición debe cumplir el vector de producción para que los ingresos sean iguales a los egresos en cada sector.

(b) Calcular el vector de producción si el sector C fija su producción en 70 unidades.

64 En una economía sencilla la distribución insumo-producto se lleva de acuerdo con la siguiente tabla (hipotética):

Consumidor	Productor		
	A	B	C
A	2/7	3/14	3/7
B	3/7	4/7	3/7
C	2/7	3/14	1/7

(a) Encontrar qué condición debe cumplir el vector de producción para que los ingresos sean iguales a los egresos en cada sector.

(b) Calcular el vector de producción si el sector C fija su producción en 35 unidades.

65 En una economía sencilla la distribución insumo-producto se lleva de acuerdo con la siguiente tabla (hipotética):

Consumidor	Productor		
	A	B	C
A	1/5	2/5	2/5
B	2/5	1/5	2/5
C	2/5	2/5	1/5

(a) Encontrar qué condición debe cumplir el vector de producción para que los ingresos sean iguales a los egresos en cada sector.

(b) Calcular el vector de producción si el sector C fija su producción en 50 unidades.

66 Una economía simple (hipotética) consta de tres sectores: carbón, electricidad y acero. Todo lo que produce cada sector es consumido entre ellos mismos de acuerdo con la siguiente tabla:

Consumidor	**Productor** *Carbón*	*Electricidad*	*Acero*
Carbón	30 %	20 %	50 %
Electricidad	40 %	50 %	30 %
Acero	30 %	30 %	20 %

Encontrar los precios de equilibrio; es decir, la producción que debe tener cada sector de tal manera que sus ingresos igualen a sus egresos.

Inversa de la matriz de Leontief y método de aproximación

En los ejercicios 67 a 69 encontrar el menor entero ν que se requiere para aproximar la inversa de la matriz de Leontief para la matriz de insumos A con la tolerancia T y entonces calcular la aproximación $(I_n - A)^{-1} \approx 1 + A + \cdots + A^{\nu}$.

67 $T = 10^{-4}$ y

$$A = \begin{bmatrix} 0.01 & 0.01 & 0.02 & 0.04 \\ 0.02 & 0.01 & 0.01 & 0.01 \\ 0.02 & 0.02 & 0.00 & 0.01 \\ 0.01 & 0.02 & 0.02 & 0.00 \end{bmatrix}.$$

68 $T = 10^{-4}$

$$A = \begin{bmatrix} 0 & 0.06 & 0 & 0 & 0 \\ 0.07 & 0 & 0.03 & 0 & 0 \\ 0.01 & 0.02 & 0.05 & 0 & 0 \\ 0 & 0 & 0 & 0.05 & 0.02 \\ 0 & 0 & 0 & 0.03 & 0.04 \end{bmatrix}.$$

69 $T = 0.001$ y

$$A = \begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 \end{bmatrix}.$$

Programación lineal (respuestas en páginas 1094-1095)

Enfoque geométrico

En los ejercicios 70 a 78 encontrar el máximo y el mínimo, si existen, de la función objetivo f con las restricciones dadas; bosquejando geométricamente el conjunto factible, calculando los puntos esquina, resolviendo los sistemas lineales correspondientes y aplicando el teorema fundamental de la programación lineal.

70 Optimizar $f = x + y$

sujeto a
$$\begin{cases} 3x + y \geq 10 \\ x + 4y \leq 18 \\ 2x - 3y \leq 3 \end{cases}$$
$x, y \geq 0$

71 Optimizar $f = 4x + 2y$

sujeto a
$$\begin{cases} 3x - 2y \geq -1 \\ x + 3y \leq 18 \\ 2x + y \leq 16 \\ x - 3y \leq 1 \\ x + 3y \geq 7 \end{cases}$$
$x, y \geq 0$

72 Optimizar $f = 30x + 20y$

sujeto a
$$\begin{cases} -20x + 30y \leq 700 \\ 20x + 20y \leq 1800 \\ 20x - 20y \leq 600 \\ 20x + 30y \geq 1100 \end{cases}$$
$x, y \geq 0$

73 Optimizar $f = x + y$

sujeto a
$$\begin{cases} 3x + y \geq 11 \\ -x + 4y \leq 18 \\ 3x + 2y \leq 30 \\ x - y \leq 5 \\ x + 3y \geq 9 \end{cases}$$
$x, y \geq 0$

74 Optimizar $f = 3x + 4y$

sujeto a
$$\begin{cases} 5x + 5y \geq 25 \\ 4x + 8y \geq 32 \\ 3x + 6y \geq 18 \end{cases}$$
$x, y \geq 0$

75 Optimizar $f = 2x + 4y$

sujeto a
$$\begin{cases} 3x + y \geq 9 \\ 2x + 2y \geq 10 \\ 3x - 4y \leq 8 \\ 3x - y \leq 20 \end{cases}$$
$x, y \geq 0$

76 Optimizar $f = -4x + 2y$

sujeto a
$$\begin{cases} x - y \geq 1 \\ x - 6y \leq -9 \end{cases}$$
$x, y \geq 0$

77 Las mismas restricciones que el problema anterior, pero la función objetivo $f = 4x + 2y$.

78 Las mismas restricciones del ejercicio 76, pero con la función objetivo $f = 2x - 3y$.

79 Una corredora de inversiones desea invertir hasta 25 000 dólares. Puede comprar una obligación de tipo T, que produce 12 % de rendimiento sobre la cantidad invertida, y puede comprar una de tipo S, que produce 18 % de rendimiento sobre la cantidad invertida. Desea invertir en la obligación tipo T al menos tanto como en la obligación del tipo S; invertir al menos 5 000 dólares en la obligación tipo T y no más de 8 000 dólares en la obligación tipo S. ¿Cúanto debe invertir en cada tipo de obligación para maximizar su rendimiento?

80 Un dietólogo quiere usar el alimento 1 y el alimento 2 para proporcionar ciertas cantidades mínimas de vitaminas A, B y C. Los datos para las cantidades de vitaminas, las necesidades y los costos se dan en la tabla siguiente. ¿Qué cantidad de cada alimento deberá usarse para satisfacer las necesidades vitamínicas de la manera más económica?

Vitamina	Alimento 1	Alimento 2	Requerimientos
A	1 unidad/g	2 unidades/g	120 unidades
B	3 unidades/g	1 unidad/g	150 unidades
C	3 unidades/g	3 unidades/g	240 unidades
Costo	10 centavos/g	15 centavos/g	

81 Resolver el problema anterior si

(a) El costo por gramo del alimento 1 es de 20 centavos y el del alimento 2 es de 15 centavos por gramo.
(b) El costo por gramo del alimento 1 es de 25 centavos y el del alimento 2 es de 20 centavos por gramo.
(c) El costo por gramo del alimento 1 es de 10 centavos y el del alimento 2 es de 30 centavos por gramo.

82 Sea r la razón del costo del alimento 1 al alimento 2 del problema del dietólogo del ejercicio 80.

(a) ¿Cómo debe ser el valor de r para que sea más económico comprar sólo el alimento 1?
(b) ¿Cómo debe ser el valor de r para que sea más económico comprar sólo el alimento 2?

83 Una expendedora de venta de tabaco al mayoreo recibe mensualmente 300 kg de tabaco del tipo R y 360 kg de tabaco del tipo S. Puede vender una mezcla A, que consta de 3 partes de tabaco tipo R por 2 partes de S, con una ganancia de \$120 por kg; y la mezcla B, que consta de 2 partes de tabaco tipo R por 3 partes de tabaco tipo S, con una ganancia de \$140 por kg. Hallar la cantidad de cada mezcla que la casa deberá preparar mensualmente para maximizar la ganancia por la venta de estas dos mezclas.

84 Resolver el problema anterior si:

(a) La ganancia por la mezcla A es de \$170 por kg y por la mezcla B de \$100 por kg.

(b) La ganancia por la mezcla A es de \$40 por kg y por la mezcla B de \$80 por kg.

85 Un jardinero puede comprar dos tipos de fertilizantes, el fertilizante A y el B. El fertilizante A contiene 10% de nitrógeno, 10% de fósforo, 10% de potasio y 5% de calcio; mientras que el fertilizante B contiene 5% de nitrógeno, 6% de fósforo, 10% de potasio y 10% de calcio. El fertilizante A cuesta \$10 el kg y el fertilizante B \$15 el kilogramo. El jardinero desea poner al menos 2 kg de nitrógeno, 2 kg de fósforo, 3 kg de potasio y 2 kg de calcio en el jardín. ¿Cúantos kilos de cada tipo de fertilizante deberá mezclar para fertilizar de la manera más económica?

86 Resolver el problema anterior si los costos por kilogramo de los fertilizantes A y B son \$15 y \$10, respectivamente.

87 Resolver el ejercicio 85 si los costos por kilogramo de los fertilizantes A y B son \$8 y \$18, respectivamente.

Método simplex para el problema estándar de programación lineal

Resolver los ejercicios 88 a 97 aplicando el método simplex para el problema estándar de programación lineal.

88 Maximizar $P = 5x_1 + 8x_2 - 4x_3$

sujeto a $\begin{cases} x_1 - 2x_2 + 3x_3 \le 4 \\ 2x_1 + x_2 + 2x_3 \le 3 \end{cases}$

$x_1, x_2, x_3 \ge 0$

89 Maximizar $P = 3x_1 + x_2 - 4x_3$

sujeto a $\begin{cases} x_1 + 2x_2 + x_3 \le 6 \\ x_1 + 2x_2 - x_3 \le 7 \\ 2x_1 + x_2 + 3x_3 \le 8 \end{cases}$

$x_1, x_2, x_3 \ge 0$

90 Maximizar $P = 3x_1 - 2x_2 - x_3$

sujeto a $\begin{cases} x_1 + x_2 - x_3 \le 8 \\ 2x_1 + x_2 + x_3 \le 6 \end{cases}$

$x_1, x_2, x_3 \ge 0$

91 Maximizar $P = 3x_1 + x_2 + x_3$

sujeto a $\begin{cases} x_1 + x_2 + x_3 \le 8 \\ 2x_1 + x_2 + 4x_3 \le 6 \end{cases}$

$x_1, x_2, x_3 \ge 0$

92 Maximizar $P = 2x_1 + 4x_2 + 2x_3 + x_4$

sujeto a $\begin{cases} 2x_1 + x_2 + 2x_3 + 3x_4 \le 14 \\ 2x_2 + 3x_3 + 2x_4 \le 20 \\ 2x_1 + x_2 + 4x_3 \le 16 \end{cases}$

$x_1, x_2, x_3, x_4 \ge 0$

93 Minimizar $C = x_1 - 3x_2 + 8x_3$

sujeto a $\begin{cases} -2x_1 + 2x_2 + x_3 \le 20 \\ -x_1 + 2x_2 - 2x_3 \le 4 \\ 4x_1 + 8x_2 - 3x_3 \le 16 \end{cases}$

$x_1, x_2, x_3 \ge 0$

94 Maximizar $P = x_1 + 2x_2 + 4x_3 - x_4$

sujeto a $\begin{cases} 5x_1 + 2x_3 + 6x_4 \le 30 \\ 4x_1 + 2x_2 + 2x_3 + 8x_4 \le 40 \end{cases}$

$x_1, x_2, x_3, x_4 \ge 0$

95 Minimizar $C = -2x_1 + 2x_2 + x_3 - 3x_4$

sujeto a $\begin{cases} 4x_1 + 6x_2 + 2x_3 - 2x_4 \le 8 \\ 3x_1 + x_2 - 4x_3 + 5x_4 \le 10 \end{cases}$

$x_1, x_2, x_3, x_4 \ge 0$

96 Minimizar $C = -4x_1 + 2x_2 + x_3 - 3x_4$

sujeto a $\begin{cases} 2x_1 + 4x_2 + 5x_3 + 6x_4 \le 24 \\ 4x_1 + 4x_2 + 2x_3 + 2x_4 \le 6 \end{cases}$

$x_1, x_2, x_3, x_4 \ge 0$

97 Minimizar $C = -x_1 + 3x_2 - 4x_3$

sujeto a $\begin{cases} -8x_1 + 5x_2 + 4x_3 \le 30 \\ -2x_1 + 6x_2 + x_3 \le 5 \\ -2x_1 + 2x_2 + x_3 \le 15 \end{cases}$

$x_1, x_2, x_3 \ge 0$

98 Una compañía dispone de 200 kg de cacahuates, 40 kg de almendras y 20 kg de nueces. Empaca tres tipos de latas de 400 gramos con estas semillas procesadas. La lata *I* contiene 200 gramos de cacahuates, 100 gramos de almendras y 100 gramos de nueces; la lata *II* contiene 300 gramos de cacahuates y 100 gramos de almendras; la lata *III* contiene 400 gramos de cacahuates. Las latas *I*, *II* y *III* se venden en $80, $60 y $40, respectivamente. ¿Cuántas latas de cada tipo debe empacar la compañía para maximizar sus ingresos?

99 Un fabricante de muebles manufactura gabinetes para televisión, sistemas estereofónicos y gabinetes para despensa. El proceso de fabricación consta de tres etapas: armar, decorar y embalar. La siguiente tabla contiene los tiempos que se utilizan en el proceso para cada artículo, en cada una de esas etapas (en horas), el número de horas disponible para cada una de ellas y la utilidad por cada unidad fabricada:

	Gabinete TV	**Gabinete ST**	**Despensero**	**Tiempo disponible**
Armar	3	10	5	10 000
Decorar	5	8	7	8000
Embalar	0.1	0.2	0.1	140
	$900	$100	$950	

¿Cuántas unidades de cada tipo de gabinete debe fabricar para maximizar su utilidad?

100 Una empresa fabrica yates en tres modelos. El modelo *I* requiere 2 semanas de fabricación, el modelo *II* necesita 2 semanas y el modelo *III* requiere de 4 semanas en su fabricación. La empresa dispone de 50 semanas para la fabricación de los yates que va a producir y no puede manufacturar más de un total de 20 de estas embarcaciones. Si las ganancias por las ventas del modelo *I*, el modelo *II* y el modelo *III* son de $400 000, $300 000 y $600 000, respectivamente, encontrar el número de cada modelo que debe fabricar para maximizar su ganancia.

101 Un granjero dispone de 50 acres de terreno para sembrar algo de soya, maíz y trigo. El costo de cultivo por acre, los días de trabajo que se requieren por acre, la utilidad por acre (en dólares) y las restricciones en tiempo y dinero se indican en la siguiente tabla :

	Maíz	**Soya**	**Trigo**	**Total disponible**
Costo del cultivo por acre	$60	$30	$50	$1 800
Días de trabajo por acre	3 días	4 días	5 días	150 días
Utilidad por acre	$250	$150	$300	

Encontrar el número de acres que se deben sembrar con cada cultivo para maximizar la ganancia.

102 Una empresa fabricante de estufas tiene tres plantas, en Monterrey, Guadalajara e Irapuato. Envía a la Ciudad de México mensualmente no más de 400 lavadoras procedentes de estas plantas. Los costos por unidad de envío de Guadalajara, Irapuato y Monterrey a la Ciudad de México son de $30, $20 y $40, respectivamente; mientras que por mano de obra tiene que gastar, en sendas ciudades, por empaque de cada unidad $6, $8 y $6. Las utilidades por cada lavadora que produce en Monterrey, Guadalajara e Irapuato son de $800, $600 y $700, respectivamente. Encontrar la cantidad de lavadoras que debe enviar de cada planta a la Ciudad de México para maximizar su utilidad, si dispone de $10 000 para los embarques y $3 000 para los empaques en total.

Restricciones generales y método simplex de dos fases

En los ejercicios 103 a 107 aplicar el método simplex de dos fases para resolver el problema de optimización dado.

103 Minimizar $C = 7x_1 + 5x_2 + 6x_3$

sujeto a $\begin{cases} x_1 + x_2 + x_3 = 10 \\ x_1 + 2x_2 + 3x_3 \leq 19 \\ 2x_1 + 3x_3 \geq 21 \end{cases}$

$x_1, x_2, x_3 \geq 0$

104 Minimizar $C = 7x_1 + 5x_2 + 6x_3$

sujeto a $\begin{cases} x_1 + x_2 + x_3 = 10 \\ x_1 + 2x_2 + 3x_3 \leq 19 \\ 2x_1 + 3x_2 \geq 21 \end{cases}$

$x_1, x_2, x_3 \geq 0$

105 Maximizar $P = 7x_1 + 5x_2 + 6x_3$

sujeto a $\begin{cases} x_1 + x_2 + x_3 = 10 \\ x_1 + 2x_2 + 3x_3 \leq 19 \\ 2x_1 + 3x_3 \geq 21 \end{cases}$

$x_1, x_2, x_3 \geq 0$

106 Maximizar $P = 2x_1 + 4x_2 - 2x_3$

sujeto a $\begin{cases} x_1 + 3x_2 + x_3 \leq 9 \\ 2x_1 + 3x_2 - x_3 \geq 16 \\ 3x_1 - 2x_2 + 2x_3 \geq 20 \end{cases}$

$x_1, x_2, x_3 \geq 0$

107 Minimizar $C = 2x_1 + x_2 + 3x_3$

sujeto a $\begin{cases} 3x_1 - x_2 - 4x_3 \leq -12 \\ x_1 + 3x_2 + 2x_3 \geq 10 \\ x_1 - x_2 + x_3 \geq 5 \end{cases}$

$x_1, x_2, x_3 \geq 0$

108 En la producción de un potente insecticida en polvo para uso agrícola, se mezclan tres sustancias químicas y se venden en bolsas de 100 kg. Los costos por kilogramo de la primera, segunda y tercera sustancia son de \$15, \$20 y \$10, respectivamente. Por razones de efectividad la mezcla debe contener por lo menos 30 kg de la primera sustancia y la cantidad de la tercera sustancia no debe exceder a la cantidad de la segunda sustancia. ¿Qué cantidad de cada sustancia debe tener la mezcla de 100 kg para satisfacer los requerimientos a un costo mínimo?

109 Reducir a un problema de dos variables el ejercicio anterior y comprobar la respuesta.

110 Una fábrica tiene dos bodegas en dos ciudades C y D con 40 y 30 unidades de un mismo artículo almacenadas, respectivamente. Se requieren enviar 30 unidades a una ciudad B_1 y 40 unidades a una ciudad B_2. Los costos unitarios de envío son los siguientes:

- de C a B_1: \$80
- de C a B_2: \$70
- de D a B_1: \$60
- de D a B_2: \$40

Encontrar el número de artículos que la empresa debe enviar de cada bodega a cada ciudad para satisfacer la demanda a un costo mínimo de envío.

111 Resolver el ejercicio anterior si los costos de envío cambian a

- de C a B_1: \$70
- de C a B_2: \$40
- de D a B_1: \$30
- de D a B_2: \$40

112 Un inversionista tiene 15 000 dólares con los cuales desea obtener el mayor rendimiento posible. Desea invertir en acciones, bonos y en una cuenta de ahorros. Las acciones le pueden dar de rendimiento 10 % de la inversión, los bonos 7 % y en la cuenta de ahorros gana 5 %. Dado que las acciones representan una inversión de alto riesgo, quiere invertir en ellas no más de la mitad de lo que va a invertir en bonos y no más de lo que deposite en la cuenta de ahorros. Además, para mayor seguridad ha decidido depositar por lo menos 3 000 dólares en la cuenta de ahorros. ¿Qué cantidad debe invertir en cada de uno de estos instrumentos financieros?

113 Resolver el ejercicio anterior si las acciones, los bonos y la cuenta de ahorro tienen rendimientos de 8 %, 7 % y 5 %, respectivamente.

Dualidad

En los ejercicios 114 a 118*(a)* establezca el dual del problema de programación lineal dado y *(b)* resuelva por el método simplex el que convenga para determinar la solución de los dos como se establece en el teorema de dualidad (6.4).

114 Minimizar $C = 3y_1 + 4y_2$

sujeto a $\begin{cases} 2y_1 + y_2 \geq 2 \\ 2y_1 + y_2 \geq 6 \end{cases}$

$y_1, y_2 \geq 0$

115 Minimizar $C = 2y_1 + 3y_2 + 4y_3$

sujeto a $\begin{cases} y_1 - 2y_2 - 3y_3 \geq -2 \\ y_1 + y_2 + y_3 \geq 2 \\ 2y_1 + y_3 \geq 3 \end{cases}$

$y_1, y_2, y_3 \geq 0$

116 Maximizar $P = 4x_1 - 2x_2$

sujeto a $\begin{cases} x_1 + 2x_2 \leq 8 \\ 3x_1 + x_2 \leq 15 \end{cases}$

$x_1, x_2 \geq 0$

117 Minimizar $C = 2y_1 + 3y_2$

sujeto a $\begin{cases} y_1 + 2y_2 \geq 6 \\ 3y_1 + 4y_2 \geq 10 \end{cases}$

$y_1, y_2 \geq 0$

118 Minimizar $C = 4y_1 + 3y_2$

sujeto a $\begin{cases} 2y_1 + y_2 \geq 6 \\ y_1 + y_2 \geq 5 \\ y_1 + 2y_2 \geq 8 \\ y_1 + 4y_2 \geq 16 \end{cases}$

$y_1, y_2 \geq 0$

119 Un paquete nutricional se forma por la combinación de tres tipos de alimentos. El alimento *I* contiene 2 gramos de proteína y 4 gramos de carbohidratos por unidad; el alimento *II* contiene 3 gramos de proteína y un gramo de carbohidratos por unidad; el alimento *III* contiene 4 gramos de proteína y 2 gramos de carbohidratos por unidad. El paquete debe contener al menos 20 gramos de proteínas y 10 gramos de carbohidratos. Si los costos por unidad del alimento *I*, *II* y *III* son \$2, \$1 y \$3, respectivamente, encontrar el número de unidades que debe contener de cada tipo de alimento el paquete para satisfacer los requerimientos alimenticios a un costo mínimo. Para ello, plantear el problema dual y resolver por el método simplex.

120 Una empresa produce láminas de fibra de vidrio en tres tipos: *A*, *B* y *C*, en dos talleres. Los costos de producción diarios de cada taller, la producción diaria, la demanda semestral y los costos diarios de operación de cada taller están contenidos en la siguiente tabla:

Tipo de lámina	Taller 1 por día	Taller 2 por día	Demanda semestral
A	200 láminas	100 láminas	7 000 láminas
B	60 láminas	20 láminas	2 200 láminas
C	40 láminas	80 láminas	4 000 láminas
Costos diarios	\$4 000	\$2 000	

Encontrar el número de días que debe trabajar cada taller para satisfacer la demanda a un costo mínimo. Para ello, establecer el problema dual y resolverlo.

Teoría de juegos (respuestas en páginas 1095-1096)

Juegos estrictamente determinados y puntos silla

En los ejercicios 121 a 126 *(a)* hallar la matriz de pagos para el problema dado; *(b)* establecer si el juego es estrictamente determinado; *(c)* en caso de que el juego sea estrictamente determinado, encontrar el valor del juego; esto es, el punto silla, por medio del criterio **minimax**; es decir, el elemento máximo de la columna y el mínimo de la fila e indicar entonces cuál es la mejor estrategia de cada jugador.

121 Dos personas, *R* y *T*, muestran simultáneamente uno o dos dedos. Si el número total de dedos es par, *R* paga a *T* la cantidad, en pesos, de los dedos mostrados; si es impar, *T* paga a *R* la cantidad, en pesos, de los dedos mostrados.

122 Dos personas, R y T, muestran simultáneamente uno o dos dedos. Si muestran el mismo número de dedos, R paga a T un peso; si el número es diferente, T paga a R un peso.

123 R y T juegan el mismo juego del ejercicio 121, pero ahora cada jugador muestra uno, dos o tres dedos.

124 Dos empresas, I y II, planean establecer por separado una sucursal en cierta ciudad en el centro o en uno de sus grandes suburbios. Si ambas establecen en el centro, I tendrá una ganancia anual de 15 000 dólares más que II; si ambas establecen sucursales en el suburbio, II tendrá una ganancia anual de 20 000 dólares más que I; si I elige el suburbio y II la ciudad, I tendrá una ganancia anual de 30 000 dólares más que II; si I pone su sucursal en la ciudad y II en el suburbio, entonces II tendrá una ganancia anual de 30 000 dólares más que I.

125 En una pequeña colonia hay dos restaurantes que compiten entre sí. El restaurante X ha concluido que si aumenta sus precios, perderá 2 % de su clientela si el restaurante Y también aumenta sus precios, perderá 4 % si Y no aumenta sus precios y perderá 10 % si Y los disminuye. Si X no hace cambios, gana 6 % si Y aumenta, pierde 4 % si Y disminuye precios y no gana ni pierde clientela si Y no hace cambios en sus precios. Si X disminuye sus precios, gana 10 % si Y los aumenta, gana 3 % si Y no hace cambios y pierde 2 % si Y disminuye precios.

126 Una planta química produce cierto tipo de material mediante cinco procesos: I, II, III, IV y V; la materia utilizada varía en contenido de un químico y puede tener 2 %, 4 %, 5 % o 6 % de éste. Con la materia que tiene 2 % del químico los procesos producen, respectivamente, 50, 50, 60, 50 y 80 toneladas del material; con la materia que contiene 4 %, los procesos producen 65, 70, 75, 90 y 75 toneladas, respectivamente; con la materia que tiene 5 % los procesos producen 30, 55, 60, 45 y 70 toneladas, respectivamente; y con la materia que contiene 6 % del químico, sendos procesos producen 40, 80, 80, 65 y 80 toneladas del material. ¿Existe un proceso y un tipo de materia prima que la planta pueda utilizar para maximizar la mínima producción esperada?

En los ejercicios 127 a 133 determinar si el juego matricial, dado por la matriz de pagos, es o no estrictamente determinado; en caso de que sea estrictamente determinado, hallar el valor del juego.

127 $\begin{bmatrix} -2 & 2 \\ -3 & 5 \end{bmatrix}$.

128 $\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$.

129 $\begin{bmatrix} 1 & -1 \\ -2 & 4 \end{bmatrix}$.

130 $\begin{bmatrix} 1 & 0 & -3 \\ -1 & 2 & 1 \\ 2 & 3 & 3 \end{bmatrix}$.

131 $\begin{bmatrix} 7 & 6 & 4 & 0 \\ -1 & 4 & 5 & -2 \\ 0 & 1 & 3 & 3 \end{bmatrix}$.

132 $\begin{bmatrix} 6 & 4 & 2 & 0 \\ -1 & 7 & 5 & -2 \\ 1 & 0 & 4 & -4 \end{bmatrix}$.

133 $\begin{bmatrix} 3 & 4 & 0 & 0 \\ -1 & 7 & 4 & -1 \\ 2 & -1 & 17 & -1 \\ 5 & 2 & -1 & 0 \end{bmatrix}$.

134 Probar que el juego con matriz de pagos $\begin{bmatrix} \alpha & \alpha \\ \beta & \gamma \end{bmatrix}$ es estrictamente determinado para todos los posibles valores de α, β y γ. Encontrar el valor del juego para cada caso.

Estrategias y pagos esperados

En los ejercicios 135 a 138 encontrar el pago esperado del juego para la matriz de pagos con las estrategias ahí dadas. El jugador P juega por filas y utiliza la estrategia $\vec{p}$; mientras que el jugador Q juega por columnas y utiliza la estrategia $\vec{q}$.

135 $\begin{bmatrix} 8 & 0 \\ -6 & 3 \end{bmatrix}$, $\vec{p} = (2/3, 1/3)$, $\vec{q} = (1/3, 2/3)$.

136 $\begin{bmatrix} 4 & -2 \\ -6 & 8 \end{bmatrix}$, $\vec{p} = (1/2, 1/2)$, $\vec{q} = (1/4, 3/4)$.

137 $\begin{bmatrix} 4 & -2 & 2 \\ -8 & 2 & -4 \end{bmatrix}$, $\vec{p} = (1/2, 1/2)$, $\vec{q} = (1/4, 1/2, 1/4)$.

138 $\begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & -2 \\ 2 & -4 & -3 \end{bmatrix}$, $\vec{p} = (1,0,0)$, $\vec{q} = (0,0,1)$.

Estrategias óptimas y valores esperados para juegos con matriz de pagos 2×2

En los ejercicios 139 a 144 utilizar las fórmulas (6.43) y (6.44) para hallar las estrategias óptimas, $\vec{p}_0$ y $\vec{q}_0$, de cada jugador y el valor esperado, $v = E(\vec{p}_0, \vec{q}_0)$, del juego matricial presentado. El jugador P juega por filas y el jugador Q juega por columnas; las estrategias óptimas de sendos jugadores son $\vec{p}_0$ y $\vec{q}_0$.

139 $\begin{bmatrix} 1 & -1 \\ -2 & 4 \end{bmatrix}$.

140 $\begin{bmatrix} 4 & 2 \\ -1 & 3 \end{bmatrix}$.

141 $\begin{bmatrix} -2 & 3 \\ 4 & -1 \end{bmatrix}$.

142 $\begin{bmatrix} 4 & 3 \\ -1 & 5 \end{bmatrix}$.

143 $\begin{bmatrix} 5 & -4 \\ -2 & 7 \end{bmatrix}$.

144 $\begin{bmatrix} -3 & 2 \\ 4 & -1 \end{bmatrix}$.

145 Encontrar las estrategias óptimas de cada jugador y el valor esperado del juego del problema del ejercicio 121.

146 Encontrar las estrategias óptimas de cada jugador y el valor esperado del juego del problema del ejercicio 122.

147 En una batalla naval, aviones bombarderos intentan hundir barcos de una flota protegida por un portaaviones con cazas. Los bombarderos pueden atacar volando alto o bajo, con mejores resultados cuando atacan a poca altura. El portaaviones puede enviar sus cazas a interceptar bombarderos a cualquier altura. Se asignan 10 puntos a los bombarderos que eviten cazas, -4 puntos a los que se enfrenten con ellos y cinco puntos extra a los bombarderos que vuelen bajo. Encontrar las estrategias óptimas para los bombarderos, los cazas y el valor esperado del juego.

148 Un narcotraficante intenta escapar de una estación del tren. El policía que lo busca desea, de ser posible, eliminarlo. Existen dos salidas de la estación; una muy concurrida (A) y otra que casi no es usada por los viajeros (B). Si el policía intercepta al narcotraficante en la salida A, únicamente lo arrestará; pero si lo encuentra en la salida B, el policía lo asesinará. Asignar un pago de 50 puntos al narcotraficante si evita al policía en la salida B y de 20 puntos si lo evita en la salida A; -100 puntos si es interceptado en la salida B y -10 puntos si es arrestado. Encontrar las estrategias óptimas de cada jugador y el valor esperado del juego.

149 Resolver el ejemplo 6.35, si la matriz de pagos es

$$\begin{bmatrix} 4 & -2 \\ 2 & 3 \end{bmatrix}.$$

Estrategias óptimas y valor esperado con programación lineal para juegos matriciales con matriz de pagos $m \times n$

En los ejercicios 150 a 153 encontrar las estrategias óptimas y el valor esperado del juego matricial utilizando el método simplex.

150 $\begin{bmatrix} 7 & 6 & 4 \\ -1 & 4 & 5 \end{bmatrix}$.

151 $\begin{bmatrix} 6 & 4 & 2 \\ -1 & 7 & 1 \\ -2 & 1 & 4 \end{bmatrix}$.

152 $\begin{bmatrix} 2 & -2 & 4 \\ 4 & 0 & 1 \\ 3 & 4 & -1 \end{bmatrix}$.

153 $\begin{bmatrix} 1 & -1 \\ 2 & -2 \\ 0 & 2 \end{bmatrix}$.

Filas y columnas dominantes o recesivas

En los ejercicios 154 a 157 encontrar las estrategias óptimas y el valor esperado del juego matricial reduciendo, si es posible, la matriz de pagos al eliminar filas y/o columnas recesivas; aplicando las fórmulas (6.43) y (6.44) o aplicando el método simplex.

154 $\begin{bmatrix} 2 & -3 & 0 \\ 3 & 1 & -2 \end{bmatrix}$.

155 $\begin{bmatrix} 8 & 4 & 2 \\ 5 & 3 & 6 \end{bmatrix}$.

156 $\begin{bmatrix} 6 & 3 & 2 \\ 4 & 2 & 6 \\ 3 & 0 & 1 \end{bmatrix}$.

157 $\begin{bmatrix} 4 & 1 & 0 & 5 \\ 3 & -2 & 1 & 2 \end{bmatrix}$.

158 Encontrar las estrategias óptimas y el valor esperado del problema planteado en el ejercicio 123.

159 Un juego llamado *piedra*, *papel* o *tijeras*, consiste en que dos oponentes de manera independiente y simultánea muestran el puño cerrado de una mano (*piedra*), una mano extendida (*papel*) o los dedos índice y medio de una mano extendidos (*tijeras*). Piedra vence a tijeras, tijeras vence a papel y papel vence a piedra; si ambos contrincantes exhiben la misma figura hay empate. Si la apuesta por partida es un peso, encontrar las estrategias óptimas de cada jugador y el valor del juego.

160 Un inversionista planea invertir 50 000 dólares en un periodo de incertidumbre: *guerra*, *rompimiento de relaciones diplomáticas* y *paz*. La inversión la puede llevar a efecto en *bonos del gobierno*, *acciones de armamento* y *acciones industriales*. La siguiente tabla contiene los rendimientos porcentuales para cada acción y eventualidad.

	Guerra	**Rompimiento**	**Paz**
Bonos	4	6	3
Acciones de armamento	20	10	−4
Acciones industriales	4	8	14

¿Qué porcentajes debe invertir en cada acción y cuál es el rendimiento esperado?

161 Tres antibióticos, A_1, A_2, A_3 y cinco tipos de bacterias, B_1, B_2, B_3, B_4, B_5, están relacionados respecto a la efectividad de los fármacos contra los bacilos, de acuerdo con la siguiente tabla de probabilidades:

	B_1	B_2	B_3	B_4	B_5
A_1	0.2	0.3	0.5	0	1
A_2	0.3	0.4	0.6	1	0
A_3	0.1	0.5	0.2	0.3	0

esto es, el antibiótico A_1 tiene una probabilidad de 0.2 de acabar con la bacteria B_1, de 0.3 de destruir a la bacteria B_2, etc. ¿En qué proporciones se deben mezclar los antibióticos para tener una mayor probabilidad de efectividad?

162 Una empresa requiere comprar 110 bombas para un proceso industrial. En el mercado existen tres calidades de bombas: calidad alta, media y baja. La bomba de calidad alta tiene un costo de $800 que cubre también una garantía total de reparación en caso de falla por defecto de fabricación; la de calidad media tiene un costo de $600 que cubre un seguro de reparación por falla de fabricación con un costo de $200 deducible. La bomba de calidad baja cuesta $300 e incluye un seguro de reparación en caso de falla de origen, con un costo de $300 por deducible. ¿Qué cantidades de cada tipo de bomba debe comprar la empresa para minimizar el valor esperado para el costo?

Cadenas de Markov (respuestas en página 1096)

163 En 2004, 43 % de los automovilistas de cierta área metropolitana tenían su automóvil asegurado en la compañía IGN, 32 % con la compañía Guardián del camino y 25 % con la compañía PGN. Para 2005, 90 % de los que estaban asegurados con IGN continuaron con esa compañía y 8 % cambiaron a la compañía Guardián del camino y 2 % a PGN; 88 % de los que estaban asegurados en la compañía Guardián del camino continuaron en ella, 10 % cambiaron a IGN y 2 % cambiaron a PGN; 90 % de los que estaban en la compañía PGN continuaron con ella, 6 % cambiaron a IGN y 4 % cambiaron a la compañía Guardián del camino.

(a) Calcular los procentajes de asegurados por cada compañía en el 2005.

(b) Si las tendencias no cambiaron para el 2006, calcular los porcentajes de asegurados por cada compañía para el 2006.

164 Si las compañías aseguradoras del ejercicio anterior no han realizado ningún cambio en sus estrategias de mercadotecnia, ¿qué proporciones de los asegurados tendrán cada una de ellas a largo plazo?

165 Suponer que las probabilidades de que una madre rubia tenga una hija rubia, con pelo oscuro o pelirroja son 0.6, 0.3 y 0.1, respectivamente; las probabilidades que una madre con pelo oscuro tenga una hija con pelo rubio, pelo oscuro o pelirroja son 0.1, 0.8 y 0.1, respectivamente; las probabilidades de que una madre pelirroja tenga una hija pelirroja, con pelo rubio o pelo oscuro son 0.5, 0.3 y 0.2, respectivamente.

(a) ¿Cuál es la probabilidad de que una abuela rubia tenga una nieta con pelo rojo?

(b) Si la distribución actual es de 65 % de mujeres con pelo oscuro, 25 % de mujeres con pelo rubio y 10 % de mujeres pelirrojas, ¿cuál será la distribución de las nietas?

(c) Calcular la distribución a largo plazo.

166 En cierta comunidad se han contrastado el nivel máximo de estudios de los hijos respecto al nivel máximo de estudios de los padres. La información se resume en la siguiente tabla:

	Padres		
Hijos	Nivel superior	Nivel medio	Nivel básico
Nivel superior	75 %	50 %	5 %
Nivel medio	20 %	30 %	30 %
Nivel básico	5 %	20 %	65 %

Así, por ejemplo, 75 % de los hijos que tienen padres con nivel superior, llegan al nivel superior, 20 % se quedan en el nivel medio y 5 % únicamente concluye el nivel básico.

(a) Si en la actualidad 15 % de la población tiene nivel máximo de estudios superiores, 35 % nivel medio y 50 % nivel básico, ¿cuál será la distribución de los nietos de esta generación?

(b) ¿A la larga cuál será la distribución de las generaciones de esta comunidad?

167 En cierta comunidad, de los hijos de los miembros del partido *PUM*, 75 % vota por el *PUM* y el resto por el *PUP*; de los hijos de los miembros del *PUP*, 65 % votan por el *PUP*, 30 % por el *PUM* y 5 % por *PAP*; de los hijos de los miembros del *PAP*, 50 % vota por el *PAP*, 25 % por el *PUP* y 25 % por el *PUM*. Si las tendencias persisten, determinar la distribución del padrón electoral que tendrá esta comunidad a largo plazo.

El método dado para encontrar el vector de distribución del estado estacionario de un proceso de Markov se basa en el hecho de que la matriz de transición sea regular (cfr. teorema 6.6 y discusión posterior a él). Sin embargo, si la matriz no es regular, se puede encontrar el vector de distribución del estado estacionario, $\vec{u}$, empleando diagonalización de matrices (*para resolver estos problemas el lector necesita*

de los conocimientos de valores propios, vectores propios y diagonalización de los apartados 5.3.1 y 5.3.2 de este libro). Para ello, si C y $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ es una diagonalización de la matriz de transición, T, del proceso, entonces $D = C^{-1}TC$ y, por tanto, $T^m = CD^mC^{-1}$ para todo $m \in \mathbb{N}$. Entonces

$$\lim_{m\to\infty} T^m = C \lim_{m\to\infty} D^m C^{-1} \tag{6.87}$$

Luego el vector de estado estacionario, de (6.51) del teorema 6.6, está dado por

$$\begin{aligned} \vec{u} &= \lim_{m\to\infty} T^m \vec{p}_0 \\ &= \left(C \lim_{m\to\infty} D^m C^{-1} \right) \vec{p}_0 \end{aligned}$$

si el límite existe. Utilizar este hecho para resolver los ejercicios 168 a 172.

168 Demostrar que la matriz de transición

$$T = \begin{bmatrix} 0.5 & 0 & 0 \\ 0.5 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

no es regular. Encontrar, si existe, el vector de estado estacionario del proceso para la distribución $\vec{p}_0 = [\ 0.3 \quad 0.5 \quad 0.2\ ]$; es decir, calcular

$$\vec{u} = \lim_{m\to\infty} T^m \vec{p}_0$$

si es que existe.

169 Sea la matriz de transición

$$T = \begin{bmatrix} 1/3 & 0 & 0 \\ 1/3 & 1 & 0 \\ 1/3 & 0 & 1 \end{bmatrix}.$$

(a) Probar que T no es regular.

(b) Calcular $\lim_{m\to\infty} T^m = L$, si existe, utilizando (6.87). $\begin{bmatrix} 0 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/2 & 0 & 1 \end{bmatrix}$

(c) Calcular, si existe

(i) $\lim_{m\to\infty} T^m \vec{p}_{01}$, si $\vec{p}_0 = [\ 0.2 \quad 0.1 \quad 0.5\ ]$. (ii) $\lim_{m\to\infty} T^m \vec{p}_{02}$, si $\vec{p}_{01} = [\ 0.3 \quad 0.4 \quad 0.3\ ]$.

(d) ¿Son iguales las columnas de L? Si la respuesta es negativa explicar por qué sucede así consultando el teorema 6.6.

(e) ¿Depende $\vec{u} = \lim_{m\to\infty} T^m \vec{p}_0$ de la distribución inicial $\vec{p}_0$? Si la respuesta es positiva explicar por qué sucede así consultando el teorema 6.6.

170 Resolver los incisos *a)*, *b)* y *c)* del ejercicio anterior para la matriz de transición

$$T = \begin{bmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 1/4 & 1/4 & 1 & 0 \\ 1/4 & 1/4 & 0 & 1 \end{bmatrix}$$

si $\vec{p}_{01} = [\ 0.1 \quad 0.3 \quad 0.2 \quad 0.4\ ]$, $\vec{p}_{02} = [\ 0.3 \quad 0.2 \quad 0.2 \quad 0.3\ ]$.

171 En una unidad de traumatología se ha determinado que al momento de llegar al hospital 35 % de sus pacientes es ambulatorio y el resto debe ser internado. Después de un mes, de los pacientes ambulatorios, 65 % se ha recuperado, 20 % permanece ambulatorio y 15 % se tiene que internar; de los pacientes internados 30 % se ha recuperado, 20 % pasa a ser ambulatorio, 20 % permanece internado y 30 % ha muerto.

(a) Determinar el porcentaje de pacientes que se recuperaron, son ambulatorios, están internados y han muerto después de un mes de su llegada.
(b) Determinar los procentajes que se tendrán de los pacientes de este hospital que se reuperarán definitivamente y los que morirán si se conservan los porcentajes de transición a largo plazo.

172 En cierto municipio existe un sistema intensivo en dos niveles, de un año cada uno, para terminar la educación secundaria; abierto a todo público que tenga certificado de educación primaria. De los estudiantes que están inscritos en el primer año, 10 % repite este nivel, mientras que 75 % pasa al siguiente nivel y 15 % se da de baja. De los estudiantes que están en el segundo año, 10 % repite el segundo nivel, 80 % se gradúa y 10 % se da de baja. Si en el otoño de 2008 la escuela tenía matriculados a 60 % de los alumnos en el primer nivel y 40 % de los estudiantes en el segundo nivel. Hallar los porcentajes de los estudiantes que se gradúan y los que se dan de baja de este plan de estudios, si se conservan los porcentajes de transición a largo plazo.

Sistemas lineales de ecuaciones diferenciales (respuestas en páginas 1096-1097)

En los ejercicios 173 a 180 resolver el sistema lineal de ecuaciones diferenciales presentado.

173 $$\begin{aligned}\dot{x}_1 &= 3x_1 - 2x_2\\ \dot{x}_2 &= 4x_1 - 3x_2\end{aligned}$$

174 $$\begin{aligned}\dot{x}_1 &= 5x_1 - 4x_2\\ \dot{x}_2 &= 8x_1 - 7x_2\end{aligned}$$

175 $$\begin{aligned}\dot{x}_1 &= -5x_1 + 12x_2\\ \dot{x}_2 &= -2x_1 + 5x_2\end{aligned}$$

176 $$\begin{aligned}\dot{x}_1 &= -3x_1 + 2x_2\\ \dot{x}_2 &= -4x_1 + 3x_2\\ \dot{x}_3 &= 4x_1 - 2x_2 + x_3\end{aligned}$$

177 $$\begin{aligned}\dot{x}_1 &= 5x_1 - 3x_3\\ \dot{x}_2 &= 6x_1 - x_2 - 3x_3\\ \dot{x}_3 &= 6x_1 - 4x_3\end{aligned}$$

178 $$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} = \begin{bmatrix}6x_1 + 3x_2 - 3x_3\\ -2x_1 - x_2 + 2x_3\\ 16x_1 + 8x_2 - 7x_3\end{bmatrix}$$

179 $$\begin{aligned}\dot{x}_1 &= 6x_1 - x_2 - 2x_3\\ \dot{x}_2 &= 4x_1 + x_2 - 2x_3\\ \dot{x}_3 &= 4x_1 + x_2 - 2x_3\end{aligned}$$

180 $$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\end{bmatrix} = \begin{bmatrix}-4x_1 - 2x_3 - 3x_4\\ -2x_2\\ 7x_1 + 5x_3 + 3x_4\\ 2x_1 + 2x_3 + x_4\end{bmatrix}$$

(Separación de variables.) Una ecuación diferencial de la forma

$$\frac{dy}{dx} = f_1(x)f_2(y)$$

es de variables separables. Para resolverla:

- Se separan las variables $\dfrac{dy}{f_1(y)} = f(x)dx$
- Se integra $\displaystyle\int \frac{dy}{f_1(y)} = \int f(x)dx$

En los ejercicios 181 a 183, resolver las ecuaciones de variables separables dadas.

181 $y' = xy$. **182** $y' = -x/y$. **183** $yy' = \sqrt{y^2 - a^2}$.

Una ecuación diferencial lineal, de orden dos, homogéna y de coeficientes constantes tiene la forma

$$y'' + py' + qy = 0 \tag{6.88}$$

donde $p, q \in \mathbb{R}$. Los ejercicios 184 a 191 están diseñados con el objetivo de dar un método para hallar la **solución general** de esta ecuación; es decir, un algoritmo para encontrar un par de funciones y_1 y y_2 tales que:

- Para todo par de constantes $C_1, C_2 \in \mathbb{R}$,
  $y_H = C_1 y_1 + C_2 y_2$ es solución de (6.88).
- Cualquier solución φ de (6.88) tiene la forma
  $\varphi = C_1 y_1 + C_2 y_2$ para cierto par de constantes $C_1, C_2 \in \mathbb{R}$.

184 Probar que si y_1, y_2 son soluciones de (6.88), entonces $y = C_1 y_1 + C_2 y_2$ también es solución. Concluir que el conjunto de todas las soluciones de la ecuación diferencial (6.88) es un subespacio vectorial de $\mathfrak{F}(\mathbb{R})$, el espacio de funciones con dominio $\mathbb{R}$ y valores reales.

185 Para la ecuación diferencial (6.88), hacer los cambios

$$x_1 = y$$
$$x_2 = y'$$

y mostrar que se obtiene el sistema lineal

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -qx_1 - px_2$$

o, en forma matricial,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{6.89}$$

donde $A = \begin{bmatrix} 0 & 1 \\ -q & -p \end{bmatrix}$.

186 Encontrar el polinomio característico de la matriz A del sistema (6.89). Observe que el polinomio que resulta se puede obtener de manera formal a partir de la ecuación diferencial:

$$y'' + py' + qy \to \lambda^2 + p\lambda + q$$

187 Si los valores propios de la matriz A del sistema (6.89) —las raíces del polinomio característico— son reales y distintos, λ_1 y λ_2, encontrar una diagonalización de la matriz A para resolver el sistema y probar que

$$y = x_1 = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x}$$

es entonces solución de (6.88), donde $C_1, C_2 \in \mathbb{R}$ son cualquier par de constantes.

188 Sean I un intervalo de $\mathbb{R}$, $f : I \to \mathbb{C}$ una función de variable real con coeficientes constantes, con $f(x) = f_1(x) + i f_2(x)$ donde $f_1 : I \to \mathbb{R}$, $f_2 : I \to \mathbb{R}$ son la parte real y la parte imaginaria, respectivamente,

de f. Se dice que f es derivable en x si f_1 y f_2 son derivables en x y, en tal caso, se define $f'(x) = f_1'(x) + if_2'(x)$. Probar que $y = f(x)$ es solución de la ecuación diferencial (6.88) si y sólo si f_1 y f_2 son también soluciones de esta ecuación diferencial.

189 Si los valores propios de la matriz A del sistema (6.89) —las raíces del polinomio característico— son complejos conjugados, $\lambda_1 = \alpha + i\beta$ y $\lambda_2 = \alpha - \beta i$. Mostrar que una diagonalización de la matriz A (sobre $\mathbb{C}$) está dada por $C = \begin{bmatrix} 1 & 1 \\ \alpha + i\beta & \alpha - i\beta \end{bmatrix}$, $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. Se define, para cada $\theta \in \mathbb{R}$, $e^{i\theta} = \cos\theta + i\,\text{sen}\,\theta$, utilizar el ejercicio anterior y el ejercicio 184 para demostrar que

$$y = x_1 = C_1 e^{\alpha x}\cos\beta x + C_2 e^{\alpha x}\,\text{sen}\,\beta x$$

es solución de (6.88), donde $C_1, C_2 \in \mathbb{R}$, son cualquier par de constantes.

190 Si los valores propios de la matriz A del sistema (6.89) —las raíces del polinomio característico— son reales e iguales, $\lambda_1 = \lambda_2 = \lambda$, mostrar que $f_1(x) = e^{\lambda x}$ y $f_2(x) = xe^{\lambda x}$ son soluciones de la ecuación diferencial (6.88).

191 Por el ejercicio 184, el conjunto de soluciones de la ecuación diferencial (6.88) es un espacio vectorial, se puede probar que éste tiene dimensión 2. Probar el siguiente algoritmo para resolver la ecuación diferencial

$$y'' + py' + qy = 0$$

utilizando los ejercicios 187, 189 y 190:

(a) Se forma el polinomio característico

$$\lambda^2 + p\lambda + q$$

y se calculan sus raíces λ_1, λ_2.

(b) Si $\lambda_1, \lambda_2 \in \mathbb{R}$ y $\lambda_1 \neq \lambda_2$, entonces todas las soluciones de la ecuación diferencial son de la forma

$$y = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x} \qquad (6.90)$$

para algún par de constantes $C_1, C_2 \in \mathbb{R}$.

(c) Si $\lambda_1, \lambda_2 \in \mathbb{R}$ y $\lambda_1 = \lambda_2$, entonces todas las soluciones de la ecuación diferencial son de la forma

$$y = C_1 e^{\lambda_1 x} + C_2 x e^{\lambda_2 x} \qquad (6.91)$$

para algún par de constantes $C_1, C_2 \in \mathbb{R}$.

(d) Si $\lambda_1, \lambda_2 \in \mathbb{C}$ y $\lambda_1 = \alpha + i\beta$, $\lambda_2 = \alpha - i\beta$, entonces todas las soluciones de la ecuación diferencial son de la forma

$$y = C_1 e^{\alpha x}\cos\beta x + C_2 e^{\alpha x}\,\text{sen}\,\beta x \qquad (6.92)$$

para algún par de constantes $C_1, C_2 \in \mathbb{R}$.

En los ejercicios 192 a 197 utilizar el método del ejercicio anterior para resolver las ecuaciones diferenciales dadas.

192 $y''+2y'-3y=0.$ **193** $y''-2y'-8y=0.$ ***194*** $y''-8y'+16y=0.$

195 $y''+2y'+y=0.$ ***196*** $y''-4y'+13y=0.$ **197** $y''-2y'+2y=0.$

Una ecuación diferencial lineal no homogénea con coeficientes y término independiente constantes, tiene la forma

$$y''+py'+qy=A \tag{6.93}$$

donde $p,q,A\in\mathbb{R}$ son constantes dadas. Los ejercicios 198 a 200 contienen un método para resolver este tipo de ecuaciones.

198 Si $q\neq 0$, mostrar:

(a) La función constante $y=p_0=A/q$ es solución de la ecuación diferencial (6.93).

(b) Sea $y_H=C_1y_1+C_2y_2$ la solución general de la ecuación diferencial homogénea asociada

$$y''+py'+qy=0$$

donde y_H tiene una de las formas (6.90) o (6.91) o (6.92). Mostrar que toda solución de (6.93) tiene la forma

$$y=y_H+p_0$$

para cierto par de constantes $C_1,C_2\in\mathbb{R}$.

199 Si $q=0$ y $p\neq 0$, proceder de manera análoga al ejercicio anterior con $p_0(x)=(A/p)x$, para mostrar que toda solución de (6.93) tiene la forma

$$y=y_H+p_0(x)$$

para ciertas $C_1,C_2\in\mathbb{R}$.

200 Si $p=q=0$, se integra dos veces.

Utilizar los ejercicios precedentes para resolver los ejercicios 201 y 202.

201 $y''-8y'+16y=5.$ ***202*** $y''-2y'=10.$

203 Resolver la ecuación diferencial no lineal de segundo orden

$$yy''+(y')^2=0$$

haciendo el cambio de variable $y'=z$ y, por tanto, utilizando la regla de cadena, $y''=zdz/dy$, para reducir a una ecuación de orden uno de variables separables (cfr. explicación que precede al ejercicio 181).

204 Sea A una matriz cuadrada de orden n que es diagonalizable y $\vec{b}\in\mathbb{R}^n$. Un **sistema afín** de ecuaciones diferenciales tiene la forma

$$\dot{\vec{x}}=A\vec{x}+\vec{b}.$$

Sean $D=\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$, C una diagonalización para la matriz A y

$$\vec{\varphi}=C[\ C_1e^{\lambda_1t}\quad C_2e^{\lambda_2t}\quad\cdots\quad C_ne^{\lambda_nt}\ ]^t$$

la solución (general) del sistema lineal asociado $\dot{\vec{x}} = A\vec{x}$. Probar que si $\vec{p}$ es una solución del sistema lineal de ecuaciones $A\vec{x} = \vec{b}$, entonces la solución del sistema afín tiene la forma

$$\vec{\psi} = \vec{\varphi} - \vec{p}.$$

Utilizar el ejercicio precedente para resolver los sistemas afines de los ejercicios 205 a 207.

205 $\begin{aligned} \dot{x}_1 &= 3x_1 - 2x_2 - 1 \\ \dot{x}_2 &= 4x_1 - 3x_2 + 2 \end{aligned}$

206 $\begin{aligned} \dot{x}_1 &= 5x_1 - 4x_2 + 1 \\ \dot{x}_2 &= 8x_1 - 7x_2 - 4 \end{aligned}$

207 $\begin{aligned} \dot{x}_1 &= -3x_1 + 2x_2 - 1 \\ \dot{x}_2 &= -4x_1 + 3x_2 + 1 \\ \dot{x}_3 &= 4x_1 - 2x_2 + x_3 - 2 \end{aligned}$

Redes eléctricas

La caída de potencial en una resistencia eléctrica está gobernada por la la ley de Ohm: $E_R = RI$ (cfr. apartado 6.2.2). En un inductor y en un capacitor se tienen las siguientes reglas, para la caída de potencial entre sus bornes, contenidas en la figura 6-36, donde L es la inductancia del inductor, medida en henrios (H), I es la corriente que circula a través de él, C es la capacitancia del capacitor, medida en faradios (F) y Q es la carga del capacitor, medida en culombios.

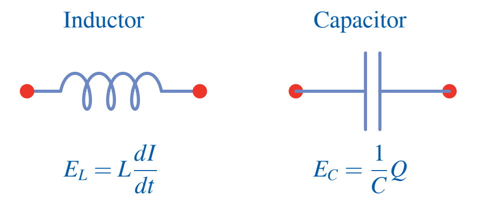


Figura 6-36 •

Considere el circuito contenido en la figura 6-37. Aplicando las leyes de Kirchhoff (cfr. sección 6.2), a la malla *abcda* se tiene

$$R_1 I + L_1 \frac{dI_1}{dt} = V;$$

pero $I = I_1 + I_2$, por tanto,

$$R_1 I_1 + R_1 I_2 + L_1 \frac{dI_1}{dt} = V \tag{6.94}$$

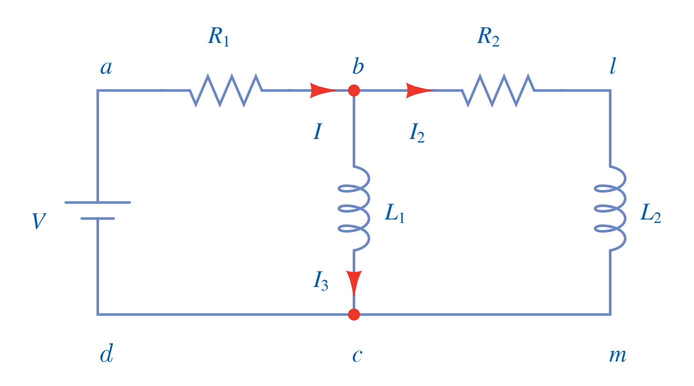


Figura 6-37 •

Aplicando ahora las leyes de Kirchhoff a la malla *ablmcda*,

$$R_1 I + R_2 I_2 + L_2 \frac{dI_2}{dt} = V$$

y como $I = I_1 + I_2$, la anterior ecuación se transforma en

$$(R_1 + R_2)I_2 + R_1 I_1 + L_2 \frac{dI_2}{dt} = V \tag{6.95}$$

Así, de (6.94) y (6.95) se obtiene el sistema afín de ecuaciones diferenciales

$$\begin{aligned} \frac{dI_1}{dt} &= -\frac{R_1}{L_1} I_1 - \frac{R_1}{L_1} I_2 + \frac{V}{L_1} \\ \frac{dI_2}{dt} &= -\frac{R_1}{L_2} I_1 - \frac{R_1 + R_2}{L_2} I_2 + \frac{V}{L_2} \end{aligned} \tag{6.96}$$

Si el circuito se conecta en $t = 0$, $I_1(0) = 0$ y $I_2(0) = 0$, las corrientes I, I_1, I_2, como funciones del tiempo, se pueden calcular resolviendo el sistema afín (6.96).

208 Encontrar las corrientes de la red eléctrica de la figura 6-37 si $R_1 = 10\,\Omega$, $R_2 = 20\,\Omega$, $L_1 = 0.02\,\text{H}$, $L_2 = 0.04\,\text{H}$ y $V = 100\,\text{V}$. ¿A qué valores se aproximan éstas para t grande?

209 En el instante $t = 0$ se cierra el interruptor de la red eléctrica mostrada en la figura 6-38; encontrar las corrientes I_1, I_2 e I_3 en función del tiempo t. ¿A qué valores se aproximan éstas para t grande?

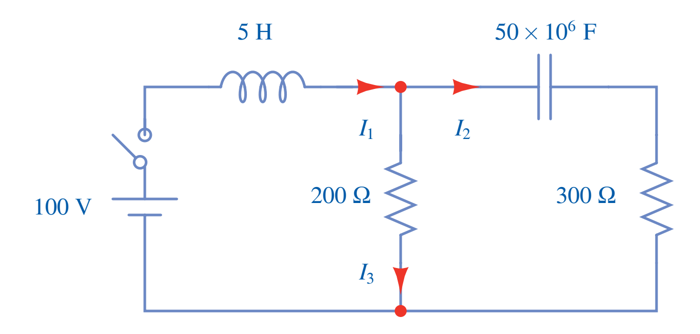


Figura 6-38 •

210 En el instante $t = 0$ se cierra el interruptor de la red eléctrica mostrada en la figura 6-39; encontrar las corrientes I_1, I_2 e I_3 en función del tiempo t. ¿A qué valores se aproximan éstas para t grande?

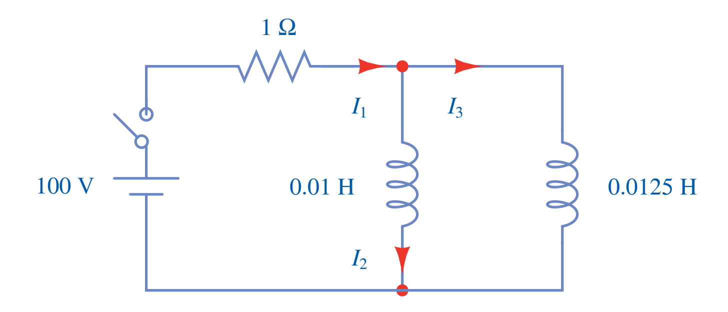


Figura 6-39 •

211 En la figura 6-20 el tanque A tiene 400 litros de agua con 600 gramos de sal disueltos y el tanque B contiene 400 litros de agua pura. A partir de $t = 0$ comienza a entrar agua pura al tanque A, a razón de 15 litros cada minuto; la salmuera es bombeada al tanque B con una velocidad de 20 litros por minuto; de este tanque la salmuera es bombeada al tanque A a razón de 5 litros cada minuto y bombeada hacia afuera a una velocidad de 15 litros por minuto. Determinar la cantidad instantánea de sal en cada tanque.

212 Resolver el problema anterior si en lugar de introducir en forma externa agua pura al tanque A, se inyecta una mezcla homogénea de salmuera que contiene 3 gramos de sal por litro. ¿Qué cantidad de sal hay, a la larga, en cada tanque?

213 En la figura 6-20 el tanque A tiene 400 litros de agua con 400 gramos de sal disuelta y el tanque B contiene 400 litros de agua con 200 gramos de sal disuelta. A partir de $t = 0$ comienza a entrar salmuera, que contiene 10 gramos de sal por litro, al tanque A, a razón de 12 litros cada minuto; la salmuera es bombeada al tanque B con una velocidad de 16 litros por minuto; de este tanque la salmuera es bombeada al tanque A a razón de 4 litros cada minuto y bombeada hacia afuera a una velocidad de 12 litros por minuto. Determinar la cantidad instantánea de sal en cada tanque. ¿Qué cantidad de sal hay, a la larga, en cada tanque?

214 Sean $\varphi(t) = \varphi_1(t) + i\varphi_2(t)$, $\psi(t) = \psi_1(t) + i\psi_2(t)$ funciones de variable real con valores complejos (cfr. ejercicio 188) y A una matriz cuadrada de orden 2. Demostrar que $[\ \varphi \quad \psi\ ]^t$ son soluciones del sistema lineal $\dot{\vec{x}} = A\vec{x}$, si y sólo si $[\ \varphi_1 \quad \psi_1\ ]^t$ y $[\ \varphi_2 \quad \psi_2\ ]^t$ son también soluciones. En tal caso, mostrar que $[\ \varphi_1 + \varphi_2 \quad \psi_1 + \psi_2\ ]^t$ son soluciones.

Utilizar el ejercicio precedente para resolver los siguientes sistemas con valores propios complejos en los ejercicios 215 y 216.

215 $\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ **216** $\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

217 Para el sistema lineal de ecuaciones de segundo orden

$$\ddot{\vec{y}} = A\vec{y}$$

hacer los cambios de variable $\vec{x}_1 = \vec{y}$ y $\vec{x}_2 = \dot{\vec{y}}$, para mostrar que se obtiene el sistema

$$\dot{\vec{x}}_1 = \vec{x}_2$$
$$\dot{\vec{x}}_2 = A\vec{x}_1$$

Osciladores armónicos acoplados

Dos cuerpos de masas m_1 y m_2 son acoplados a un par de resortes como se ilustra en la figura 6-40.

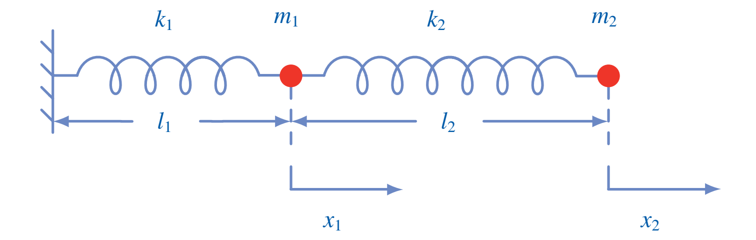


Figura 6-40 •

Los resortes tienen constantes de restitución k_1, k_2 y longitudes naturales l_1 y l_2, respectivamente. El resorte de la izquierda está sujeto a una pared. La ley de Hooke establece que la fuerza con la que actúa un resorte sobre un cuerpo que está sujeto a él es proporcional a la distancia que el resorte se ha comprimido o estirado de la posición de equilibrio; es decir, la posición natural del resorte; esto es $F = -kx$, donde $k > 0$ es la constante de restitución (o de rigidez) del resorte y x la posición relativa

al punto de equilibrio. En el instante $t = 0$ el cuerpo de masa m_1 es colocado a una distancia a de la posición de equilibrio del primer resorte y el cuerpo de masa m_2 a una distancia b de la posición de equilibrio del segundo resorte y entonces, los cuerpos son soltados para que oscilen libremente debido a la acción de los resortes. Se supone que no existen fuerzas, aparte de las de los resortes, que actúen sobre los cuerpos. Sean $x_1(t)$ y $x_2(t)$ los desplazamientos en función del tiempo para el primero y segundo cuerpo, respectivamente, relativos a las posiciones de equilibrio de cada uno de éstos (considerados positivos hacia la derecha de los puntos de equilibrio). Entonces, si en el instante t la masa m_1 ocupa la posición x_1 y la masa m_2 la posición x_2, el primer resorte actúa sobre el primer cuerpo con una fuerza $F_1 = -k_1x_1$; mientras que el segundo resorte ha experimentado una elongación (o compresión) $L_2 - (L_2 - (x_2 - x_1))$ y, por tanto, actúa sobre el primer cuerpo con una fuerza $F_2 = k_1(x_2 - x_1)$; y con una fuerza sobre el segundo cuerpo de $F_3 = -k_2(x_2 - x_1)$. Se tiene así, por la segunda ley de Newton,

$$\begin{aligned} m_1x_1'' &= -k_1x_1 + k_2(x_2 - x_1) \\ m_2x_2'' &= -k_2(x_2 - x_1) \end{aligned}$$

con las condiciones iniciales

$$\begin{aligned} x_1(0) &= a \\ x_2(0) &= b \\ x_1'(0) &= 0 \\ x_2'(0) &= 0. \end{aligned}$$

218 Si en la figura 6-40 $m_1 = 1, m_2 = 1$ (en kilogramos) , $k_1 = 2$ y $k_2 = 3$ (en newtons sobre metros), $a = -1$, $b = 1$ (en metros), hallar las posiciones $x_1(t)$ y $x_2(t)$ relativas a las posiciones de equilibrio de cada uno de los cuerpos.

219 Resolver el ejercicio precedente si además, al segundo cuerpo se le da un impulso inicial hacia la derecha de 2 m/s.

Optimización de funcionales (respuestas en página 1097)

Límites y continuidad

220 Sean $\mathbf{E}$ un espacio con producto interior $\langle\cdot,\cdot\rangle$ y $J : \mathbf{E} \to \mathbb{R}$ el funcional $J(f) = \|f\|^2$, donde $\|\cdot\|$ es la norma inducida por el producto interior; es decir, $\|f\| = \sqrt{\langle f, f\rangle}$. Demostrar que

$$\lim_{h\to 0} \frac{J(f+h) - J(f) - 2\langle f,h\rangle}{\|h\|} = 0\,.$$

221 Sea $J : C^1[0,1] \to \mathbb{R}$ el funcional definido por $J(y) = \int_0^1 (y + 2y')dx$. Probar que J es continuo en $f_0(x) = x^2$.

222 Sea $\mathbf{E}$ un espacio normado. Demostrar que $J : \mathbf{E} \to \mathbb{R}$ definido por $J(f) = \|f\|$ es continuo en todo $f \in \mathbf{E}$.

En los ejercicios 223 a 228, $\mathfrak{M}_n$ es el espacio de matrices cuadradas de orden n, con la norma $\|A\|_\infty = \text{máx}\,|a_{ij}|$. En cada caso probar que el operador J ahí definido es continuo en toda $A \in \mathfrak{M}_n$.

223 $J : \mathfrak{M}_n \to \mathfrak{M}_n\,,\, J(A) = A^2$.

224 $J : \mathfrak{M}_n \to \mathfrak{M}_n\,,\, J(A) = A + B_0$, donde B_0 es una matriz fija en $\mathfrak{M}_n$.

225 $J : \mathfrak{M}_n \to \mathfrak{M}_n$, $J(A) = A^t$.

226 $J : \mathfrak{M}_n \to \mathfrak{M}_n$, $J(A) = (A^t + B_0)^2$.

227 $J : \mathfrak{M}_n \to \mathbb{R}$, $J(A) = \text{tra}(A)$, la traza de la matriz A (la suma de los elementos de la diagonal).

228 $J : \mathfrak{M}_n \to \mathfrak{M}_n$, $J(A) = A^2 + 2A^t$.

En los ejercicios 229 a 233, $C[a,b]$ es el espacio de las funciones continuas con la norma $\|f\|_\infty = \text{máx}_{a\le x\le b} |f(x)|$. En cada caso demostrar que el operador J ahí definido es continuo en todo $f \in C[a,b]$.

229 $J : C[a,b] \to C[a,b]$, $J(f) = f_0$, donde $f_0 \in C[a,b]$ es una función fija.

230 $J : C[a,b] \to C[a,b]$, $J(f) = f$.

231 $J : C[a,b] \to C[a,b]$, $J(f) = f^2$.

232 $J : C[a,b] \to C[a,b]$, $J(f) = f^n$.

233 $J : C[a,b] \to \mathbb{R}$, $J(f) = \int_a^b f^2$

En los ejercicios 234 a 237 encuentre el conjunto en $\mathbb{R}^n$ donde la función $f : \mathbb{R}^n \to \mathbb{R}^m$ ahí definida es continua.

234 $f : \mathbb{R}^3 \to \mathbb{R}^4$, $f(x,y,z) = (-x^2 + 2yz, \frac{x-y+z}{(x^2+y^2+z^2-4)^{3/2}}, \frac{x+y}{xyz}, xyz)$.

235 $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) = (\ln(xy), \sqrt{x-y})$.

236 $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x,y,z) = \cos(x-y+z) + e^{xyz}$.

237 $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) = (x + y^2, xy + \cos(x - 3x^2 - y^2))$.

Diferenciabilidad

238 Demostrar que si **E**, **F** son espacios vectoriales normados, entonces:

(a) $J : \mathbf{E} \to \mathbf{F}$, $J(f) = f_0$, donde f_0 es un vector fijo de **F**, es diferenciable en todo punto $f \in \mathbf{E}$ y $J'(f) = 0$, la transformación cero; es decir, $J'(f)h = 0 \; \forall \, h \in \mathbf{E}$.

(b) $J : \mathbf{E} \to \mathbf{E}$, $J(f) = f$, es diferenciable en todo punto $f \in \mathbf{E}$ y $J'(f)h = h$ para todo $h \in \mathbf{E}$.

En los ejercicios 239 a 245 demostrar que el operador ahí definido es diferenciable en todo punto del conjunto G y hallar $J'(f)h$ para cada h. En algunos casos es conveniente primero encontrar la derivada direccional

$$\lim_{t\to 0} \frac{J(f+th) - J(f)}{t}$$

para conjeturar el valor de $J'(f)h$.

239 $J : \mathfrak{M}_n \to \mathfrak{M}_n$, $J(A) = A^2$, $G = \mathfrak{M}_n$.

240 $J : \mathfrak{M}_n \to \mathfrak{M}_n$, $J(A) = A^k$, donde k es un entero positivo mayor a 2, $G = \mathfrak{M}_n$.

241 $J : \mathfrak{M}_n \to \mathbb{R}$, $J(A) = \text{tra}(A)$, donde $\text{tra}(A)$ es la traza de la matriz A (la suma de los elementos de la diagonal) $G = \mathfrak{M}_n$.

242 $J : C[a,b] \to C[a,b]$, $J(f) = f^n$, donde n es un entero positivo, $G = C[a,b]$.

243 $J : C^1[a,b] \to C[a,b]$, $J(f) = (f')^n$, donde n es un entero positivo, $G = C^1[a,b]$

244 $J : G \subset C^1[a,b] \to \mathbb{R}$, $J(f) = \sqrt{1 + (f')^3}$, $G = \{f \in C^1[a,b] \,|\, f(x) \neq 0 \; \forall x \in [a,b]\}$.

245 Sea **E** un espacio con producto interior $\langle\cdot,\cdot\rangle$. Se define $J(f) = \|f\|^2$, donde $\|\cdot\|$ es la norma inducida por el producto interior; es decir, $\|f\| = \sqrt{\langle f, f\rangle}$. Demostrar que J es diferenciable en todo $f \in \mathbf{E}$ y hallar $J'(f)h$ para cada $h \in \mathbf{E}$.

Condiciones suficientes de diferenciabilidad y diferencial para funcionales en $\mathbb{R}^n$

En los ejercicios 246 a 250 demostrar que la función $f:\mathbb{R}^n\to\mathbb{R}$ es diferenciable en todo punto y utilizar la matriz jacobiana $[df(\vec{u})]$, para calcular $df(\vec{u})\vec{h}$.

246 $f:\mathbb{R}^3\to\mathbb{R}$, $f(x,y,z)=x^2-yz+xz-ze^{xy}$; $\vec{u}=(1,-1,0)$ y $\vec{h}=(1,2,1)$.

247 $f:\mathbb{R}^4\to\mathbb{R}$, $f(x_1,x_2,x_3,x_4)=x_1^2x_2x_3+x_2x_3e^{x_2x_4}-x_3x_4^2$; $\vec{u}=(1,0,-1,2)$, $\vec{h}=(1,-2,1,1)$.

248 $f:\mathbb{R}^3\to\mathbb{R}$, $f(x,y,z)=\cos(x^2+y^2-z^2)$, $\vec{u}=(0,0,\sqrt{\pi/2})$, $\vec{h}=(1,1,-\sqrt{2/\pi})$.

249 $f:\mathbb{R}^3\to\mathbb{R}$, $f(x,y,z)=\operatorname{sen}(x^2+y^2-z^2)$, $\vec{u}=(x,y,z)$, $\vec{h}=(h_1,h_2,h_3)$.

250 $f:\mathbb{R}^4\to\mathbb{R}$, $f(x_1,x_2,x_3,x_4)=\ln(x_1^2+x_2^2+x_3^2+x_4^2+1)$.

Condiciones suficientes de diferenciabilidad y diferencial para funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$

En los ejercicios 251 a 255 demostrar que la función $f:\mathbb{R}^n\to\mathbb{R}^m$ es diferenciable en todo punto y utilizar la matriz jacobiana, $[df(\vec{u})]$, para calcular $df(\vec{u})\vec{h}$.

251 $f:\mathbb{R}^3\to\mathbb{R}^3$, $f(x,y,z)=(x^2-2xy+z^2,xy+z^2,z^3y-x)$, $\vec{u}=(1,0,1)$, $\vec{h}=(1,1,1)$.

252 $f:\mathbb{R}^2\to\mathbb{R}^2$, $f(x,y)=(\ln(x^2+y^2+4),x^2+y^2)$, $\vec{u}=(1,1)$, $\vec{h}=(-3,6)$.

253 $f(x_1,x_2,x_3,x_4)=(x_1x_2x_3x_4,x_1x_2x_3,x_1x_2,x_1)$, $\vec{u}=(1,-1,1,-1)$, $\vec{h}=(1,1,1,1)$.

254 $f(x,y)=(xe^{-x^2-y^2},ye^{-x^2-y^2},e^{-x^2-y^2})$, $\vec{u}=(1,1)$, $\vec{h}=(h_1,h_2)$.

255 $f(t)=(t,\operatorname{sen}t,\cos(t))$, $\vec{u}=t$, $\vec{h}=k$.

256 Demostrar que una función $f:D\subset\mathbb{R}\to\mathbb{R}^n$ es diferenciable en $t_0\in D$ si y sólo si

$$\lim_{h\to 0}\frac{1}{h}\left(f(t_0+h)-f(t_0)\right)=f'(t_0)$$

existe y que, en tal caso, $df(t_0)h=f'(t_0)h$. Al vector $f'(t_0)$, de existir, se le llama la derivada de f en t_0.

Extremos locales y valores propios de la matriz hessiana

257 Demostrar el siguiente criterio, llamado de la segunda derivada, para funciones de dos variables reales: Sean $f:D\subset\mathbb{R}^2\to\mathbb{R}$ una función de dos variables con segundas derivadas parciales continuas en el punto crítico (x_0,y_0) y

$$D(x_0,y_0)=f_{xx}(x_0,y_0)f_{yy}(x_0,y_0)-[f_{xy}(x_0,y_0)]^2.$$

(a) Si $D(x_0,y_0)>0$ y $f_{xx}(x_0,y_0)>0$, entonces $f(x_0,y_0)$ es un mínimo relativo.
(b) Si $D(x_0,y_0)>0$ y $f_{xx}(x_0,y_0)<0$, entonces $f(x_0,y_0)$ es un máximo relativo.
(c) Si $D(x_0,y_0)<0$, entonces en (x_0,y_0) se tiene un punto silla.
(d) Si $D(x_0,y_0)=0$, el criterio no aplica para resolver la naturaleza del punto crítico.

Utilice el criterio de la segunda derivada dado en el ejercicio precedente para hallar los extremos relativos de las funciones dadas en los ejercicios 258 a 263:

258 $f(x,y)=3x^2+2xy+2x+y^2+y+4$.

259 $f(x,y)=\frac{1}{2}x^2-xy-\frac{1}{2}y^2-3x+7y$.

260 $f(x,y)=\frac{1}{2}x^4+\frac{1}{2}y^4-\frac{1}{2}x^2-xy-\frac{1}{2}y^2$.

261 $f(x,y) = xy + \frac{1}{x} + \frac{1}{y}$.

262 $f(x,y) = e^{2x+3y}(8x^2 - 6xy + 3y^2)$.

263 $f(x,y) = (5x + 7y - 25)e^{-(x^2+xy+y^2)}$.

Encontrar los extremos relativos de las funciones dadas en los ejercicios 264 a 269:

264 $f(x,y,z) = x^2 + y^2 + z^2 + 4x + 6y - 8z$.

265 $f(x,y,z) = x^2 + 2y^2 + 2z^2 - 4x + 12y + 16z$.

266 $f(x,y,z) = \frac{1}{3}x^3 + \frac{1}{2}y^2 + z^3 - xy - 3z$.

267 $f(x,y,z) = x + \frac{y^2}{4x} + \frac{z^2}{y} + \frac{2}{z}$, $x > 0, y > 0, z > 0$.

268 $f(x,y,z) = xy^2z^3(7 - x - 2y - 3z)$, $x \neq 0$, $y \neq 0$, $z \neq 0$.

269 $f(x,y,z) = \frac{1}{x} + \frac{x^2}{y} + \frac{y^2}{z} + 16z^2$

Condiciones necesarias para que cierto tipo de funcionales en espacios de dimensión infinita alcancen valores extremos

En los ejercicios 270 a 275, encontrar los puntos donde posiblemente alcanzan valores extremos los funcionales indicados con las condiciones de frontera dadas.

270 $J(f) = \int_0^1 (e^{x^2} + (f'(x))^2)dx$, $f(0) = 0$, $f(1) = 3$.

271 $J(f) = \int_{-1}^3 4(f'(x)^3 dx$, $f(-1) = 4$, $f(3) = 1$.

272 $J(f) = \int_{-1}^0 (\frac{1}{2}f(x) + f(x)f'(x) + f'(x) + \frac{1}{2}[f'(x)]^2)dx$, $f(-1) = 0$, $f(0) = 1$.

273 $J(f) = \int_0^{\pi/4} (\frac{1}{2}(f')^2 + ff' - 2f^2)$, $f(0) = 0$, $f(\pi/4) = 1$.

274 $J(f) = \int_0^1 ([f']^2 + 2f' + f^2)$, $f(0) = 1$, $f(1) = 1$.

275 $J(f) = \int_1^2 ([f']^2 + 2ff' + f^2)dx$, $f(1) = 1$, $f(2) = e$.

Cuando se requiere hallar las funciones en las que el operador integral (6.77), página 722, posiblemente alcanza un extremo sin condiciones de frontera; es decir, que satisfacen la condición de Euler (6.78), página 722, a estas funciones se les dice **extremales** del funcional (6.77), página 722. En los problemas 276 a 279 hallar los extremales del operador indicado.

276 El funcional

$$J(f) = 2\pi \int_a^b f\sqrt{1 + (f')^2}dx$$

del problema de superficie mínima de revolución.

277 El funcional

$$J(f) = \rho g \int_a^b f\sqrt{1 + (f')^2}dx$$

del problema de la catenaria.

278 $J(f) = \int_a^b \frac{\sqrt{1 + (f')^2}}{f}dx$.

279 $J(f) = \int_a^b (x^2 - (x')^2 dt$.

Dinámica de un monopolista

280 Si en el problema de dinámica de un monopolista, la función de demanda depende únicamente de p; esto es

$$y = ap + b$$

mostrar que el punto crítico para este nuevo problema es p_0, dado por (6.85), página 726.

281 En el modelo de dinámica de un monopolista demostrar que para T suficientemente grande

$$p^*(t) = p_0 + r_1 e^{-\lambda t}$$

y, por tanto, p^* tiende asintóticamente a p_0.

282 Demostrar que si p^* es el punto crítico dado por (6.86), página 726, entonces la gráfica de $p = p^*(t)$ interseca a la gráfica de $p = p_0$ a lo más una vez.

283 Sean $r_1 = p_1 - p_0$ y $r_2 = p_2 - p_0$, entonces (6.86) se puede escribir como

$$p^*(t) = \frac{r_1 - r_2 e^{\lambda T}}{1 - e^{2\lambda T}} e^{\lambda t} + \frac{r_1 - r_2 e^{-\lambda T}}{1 - e^{-2\lambda T}} e^{-\lambda t} + p_0$$

demostrar lo siguiente:

(a) Si r_1 y r_2 tienen signos opuestos, la gráfica de p^* cruza la línea $p = p_0$ una vez y no tiene tangente horizontal. Así que p^* es creciente o decreciente en todo el intervalo $[0, T]$.

(b) Si r_1 y r_2 tienen el mismo signo, entonces la gráfica de $p = p^*(t)$ no cruza la línea $p = p_0$ y tiene una tangente horizontal siempre que

$$\operatorname{sech}(\lambda T) \leq \frac{r_1}{r_2} \leq \cosh(\lambda T) \tag{6.97}$$

de otra forma no tiene tangente horizontal.

(c) Suponiendo que se cumple (6.97), mostrar:

(i) Si r_1 y r_2 son positivos, p^* tiene un mínimo. Bosquejar la gráfica de p^*.
(ii) Si r_1 y r_2 son negativos, p^* tiene un máximo. Bosquejar la gráfica de p^*.

7 Uso de tecnología

7.1 La calculadora HP 50g y álgebra lineal

La calculadora y graficadora programable HP 50g de Hewlett-Packard Company, esencialmente es una computadora de bolsillo y excelente herramienta para el cálculo simbólico y numérico en ciencias e ingeniería; en particular en álgebra lineal. Esta calculadora contiene un potente sistema algebraico computacional (CAS) que permite al usuario tener acceso a diferentes modos de operación, por ejemplo: números reales-números complejos, cálculo simbólico-cálculo numérico. La introducción y salida de información se puede ajustar de tal manera que ambos formatos estén en notación matemática, reduciendo así a un mínimo el lenguaje de intercomunicación que el usuario tiene que aprender para interactuar con la calculadora. Ésta, aparte de la portabilidad (mide 184 mm por 87 mm por 23.5 mm) es una gran ventaja que tiene la calculadora HP 50g respecto a los paquetes comerciales como Matlab y Mathematica; pues su modo de operar, en forma directamente análoga a una calculadora normal, la hace mucho más amigable al usuario que cualquier software comercial. Estas características convierten a la calculadora HP 50g en un magnífico instrumento para los estudiantes de ciencias e ingeniería y también para el ingeniero y el investigador, ya sea en el escritorio o en el campo. Naturalmente la calculadora tiene limitaciones y desventajas frente a paquetes comerciales para computadoras; las principales se presentan cuando se tienen que procesar volúmenes grandes de información y cuando la complejidad en los cálculos requiere de programación más sofisticada. En este apartado estudiaremos el uso de la calculadora HP 50g en álgebra lineal únicamente y sólo veremos los aspectos básicos. Por cuestiones de espacio, objetivos y alcance, no estudiaremos las demás herramientas que tiene esta calculadora como son sus capacidades gráficas, de programación y su uso en otras áreas de las ciencias exactas y aplicadas. Remitimos al lector a consultar el manual del usuario que acompaña a la calculadora y la dirección www.hp.com/calculators, donde encontrará material en línea, para profundizar en la aplicación de esta poderosa calculadora.

7.1.1 *Teclado y sus funciones*

La figura 7-1 contiene una imagen de la calculadora HP 50g. En ella se pueden distinguir tres partes que se utilizan para interactuar con la máquina: el teclado, la pantalla y las teclas direccionales. El teclado se emplea para introducir y obtener información de la calculadora; en la pantalla se puede desplegar la información que introduce el usuario, la información de salida que da como respuesta la calculadora a las instrucciones del operario y los menús para la interacción entre el usuario y la máquina; mientras que

Figura 7-1 • Calculadora Hp 50g.

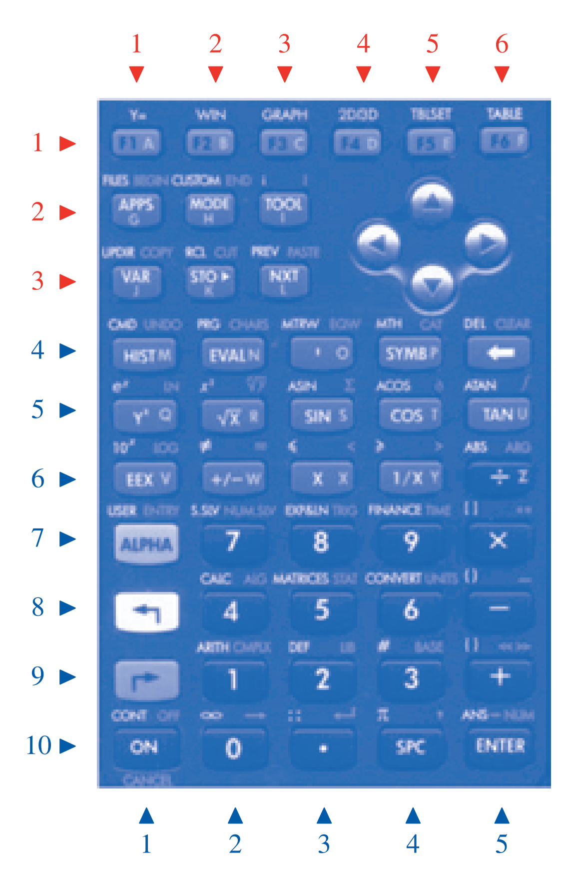


Figura 7-2 • Arreglo matricial para representar las teclas de la calculadora HP 50g.

las teclas direccionales se emplean para seleccionar opciones de los menús y desplazar el cursor a través de la pantalla.

Para poder explicar el uso del teclado hemos acomodado las teclas de la calculadora en un arreglo matricial o, mejor dicho, en dos arreglos matriciales. La figura 7-2 contiene la imagen del teclado de la calculadora HP 50g y en ella se han numerado las filas y columnas en la parte superior, inferior y en la parte lateral. Como mencionamos, en realidad se tienen dos arreglos:

- Las teclas que corresponden a las filas 1 a 3 y las columnas numeradas, en la parte superior de la imagen, del 1 al 6; en rojo ambas numeraciones para distinguir este arreglo (vea la figura 7-2).
- Las teclas que corresponden a las filas 4 a 10 y las columnas numeradas, en la parte inferior de la imagen, del 1 al 5; en azul ambas numeraciones para distinguir este arreglo (vea la figura 7-2).

Así, por ejemplo $(7, 1)$, representa la tecla que está en la fila 7 y en la columna 1, de la numeración en la parte inferior de la imagen, esto es, la tecla ALPHA. Mientras que el par $(2, 3)$ representa la tecla que se encuentra en la fila 2 y la columna 3, de la numeración en la parte superior de la imagen de la figura 7-2, es decir, TOOL. Toda tecla tiene dos o más funciones, la principal es la que está más visible al centro de la misma. Las demás funciones se encuentran en los extremos superior izquierdo (o al centro) y derecho de cada tecla resaltadas por un color en su calculadora: blanco y naranja, respectivamente; algunas teclas también contienen una letra mayúscula del alfabeto en color amarillo.

Por ejemplo, la figura 7-3 contiene la imagen de la tecla $(5, 1)$; ella tiene cinco funciones:

- Y^x (la principal).
- e^x.
- LN, la función logaritmo natural ln.
- El caracter alfabético Q mayúscula.
- El caracter alfabético Q minúscula.

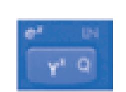


Figura 7-3 •

Para activar la función principal de cada tecla, simplemente se oprime ésta; pero para activar las demás funciones se requieren combinaciones de las siguientes teclas:

Tecla de cambio izquierdo: Al oprimir la tecla de cambio izquierdo, ↰ y posteriormente cualquier otra tecla, se activa la función que esta última contiene en el extremo superior izquierdo; es decir, la función resaltada en blanco en su calculadora.

Tecla de cambio derecho: Al oprimir la tecla de cambio derecho, y posteriormente cualquier otra tecla, se activa la función que esta última contiene en el extremo superior derecho; es decir, la función resaltada en naranja en su calculadora.

Tecla para mayúsculas: Si se oprime la tecla ALPHA y después cualquier tecla que contenga una letra mayúscula del alfabeto en color amarillo, este carácter quedará escrito en la pantalla de la calculadora.

Tecla para minúsculas: Para escribir una letra minúscula se oprime primero la tecla ALPHA seguida de la tecla cambio izquierdo: y por último la tecla que contenga la letra mayúscula que se quiere escribir en minúscula.

❍ **Nota 7.1** No hemos descrito funciones obvias que tienen algunas teclas de la calculadora HP 50g porque operan exactamente como lo hacen en las calculadoras normales. La tecla ON es la que se utiliza para encender la máquina y la tecla cambio derecho, seguida de la tecla ON, para apagarla; pues como explicamos antes, la opción OFF que aparece en color naranja en el extremo superior derecho de la tecla ON (la tecla $(10, 1)$) de su calculadora, se activa con la tecla cambio derecho (tecla $(9, 1)$).

7.1.2 La pantalla y comandos de decisión

Como explicamos antes, en la pantalla se despliegan ciertos menús que el usuario puede solicitar desde el teclado. Por ejemplo, si se oprime la tecla de cambio izquierdo $(8, 1)$ y después la tecla $(8, 3)$ (cfr. figura 7-2, pág. 762), la que tiene como función en el extremo superior izquierdo, la función (resaltada en blanco) MATRICES, en la pantalla de la HP 50g aparecerá el menú que se muestra en la figura 7-4. Entonces, para seleccionar cualquiera de los ocho submenús que contiene éste, se utilizan las flechas direccionales (cfr. figura 7-1, pág. 762) para desplazarse hacia arriba o hacia abajo en la pantalla en cada uno de los niveles, en este caso numerados del 1 al 8; note que el menú seleccionado aparece resaltado en forma sombreada en la pantalla. Una vez que se ha elegido un menú, éste se activa oprimiendo la tecla que aparece exactamente abajo de la opción OK, la tecla F6, como se ilustra en la figura 7-4.

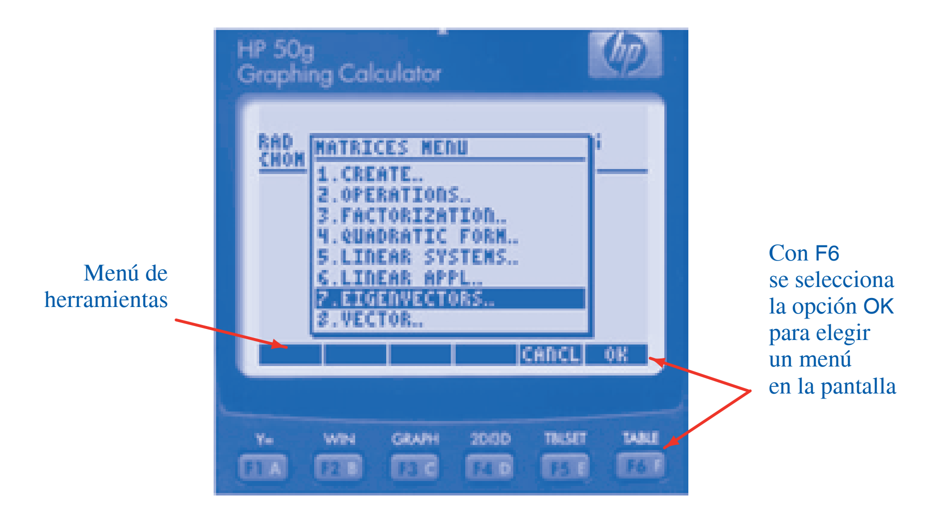


Figura 7-4 •

En general, las teclas F1 a F6 (teclas $(1,1)$ a $(1,6)$) se utilizan para activar, en cualquier estado, las opciones que estén desplegadas al pie de la pantalla, inmediatamente arriba de cada una de estas teclas. Estas opciones, que varían de acuerdo con el estado que se ha elegido, se encuentran contenidas en el llamado MENÚ DE HERRAMIENTAS mostrado, para el caso ahí señalado, en la figura 7-4. Así, si se desea cancelar el estado que muestra la pantalla de la figura 7-4, se oprime la tecla F5 para activar la opción CANCL del menú de herramientas para ese fin. Cabe señalar que un determinado estado en la pantalla puede tener más opciones en el correspondiente menú de herramientas de las que aparecen desplegadas; para hacer que la HP muestre las demás se utiliza la tecla NXT ($(3,3)$).

7.1.3 Modos de operación

La calculadora HP 50g tiene dos modos de operación, el modo ALGEBRAICO (Algebraic) y el modo NOTACIÓN POLACA INVERSA (RPN). Utilizaremos únicamente el modo algebraico, remitimos al lector familiarizado con el modo RPN y que desee usar esa opción operativa, a consultar el manual del usuario que acompaña a la calculadora. A partir de aquí, invitamos al lector a seguir simultáneamente las instrucciones que demos con su HP 50g. Una vez encendida la calculadora, seleccionamos el modo operativo con el que vamos a trabajar, en nuestro caso el algebraico. Para ello se oprime la tecla MODE, tecla $(2,2)$, y entonces se desplegará en la pantalla de la calculadora el menú mostrado en la figura 7-5*(a)*, después se debe oprimir la tecla F2 para activar la opción CHOOS del menú de herramientas. Inmediatamente la calculadora desplegará en su monitor información como la mostrada en la figura 7-5*(b)*, con las teclas direccionales hay que seleccionar la opción Algebraic y oprimir la tecla F6 para activar la opción OK y la calculadora desplegará una pantalla como la contenida en la figura 7-5*(c)*. Se oprime la tecla F3 para activar la opción CAS del menú de herramientas y se deplegará la pantalla contenida en la figura 7-5*(d)*. En este menú se pueden seleccionar las opciones que se deseen, utilizando las teclas direccionales de la calculadora y empleando las teclas F1 a F6 para activar las que convengan del menú de herramientas; nosotros, por ahora, lo dejaremos como se muestra en la pantalla de la figura 7-5*(d)*. Finalmente, hay que oprimir dos veces la tecla F6 para activar el mismo número de veces la opción OK del menú de herramientas y entonces, en el monitor de la calculadora aparecerá la información contenida en la figura 7-5*(e)*. La calculadora está ya lista para trabajar en la forma en que nos conviene para álgebra lineal.

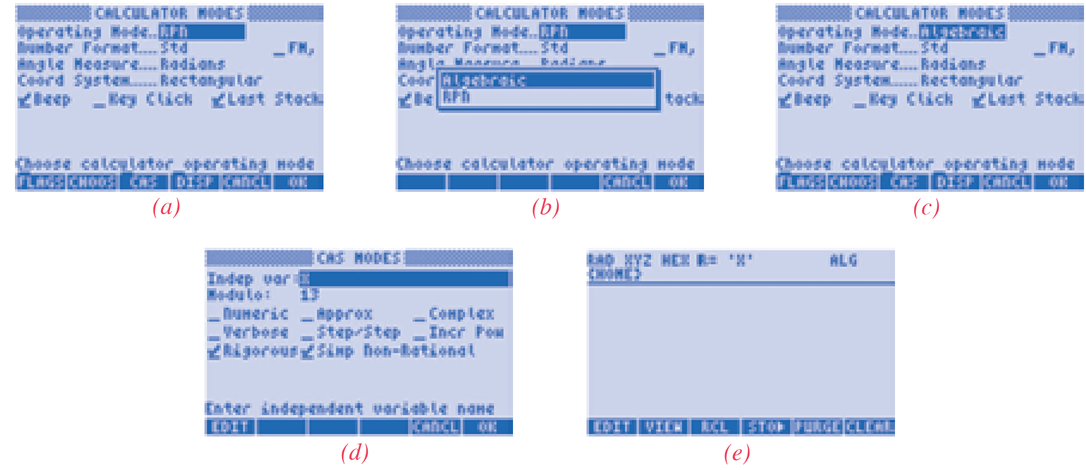


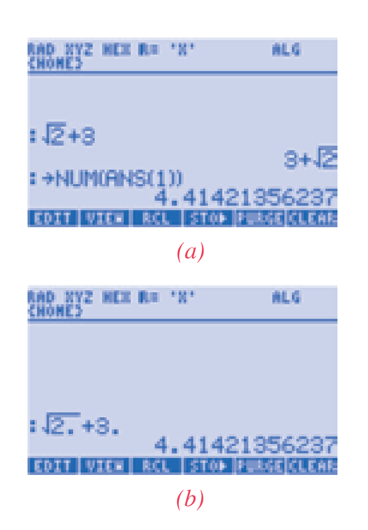
Figura 7-5 •

7.1.4 Cálculo simbólico vs. numérico y almacenamiento de objetos algebraicos

En el modo que hemos seleccionado para trabajar con la calculadora, la información que se introduce en ella y la que la máquina da como respuesta, está esencialmente en la misma notación que utilizamos en matemáticas. Por ejemplo, para calcular $\sqrt{2}+3$, sólo tenemos que oprimir sucesivamente las teclas:

- $\sqrt{\text{x}}$ (tecla $(5,2)$),
- la tecla $(9,3)$, la que contiene al número 2 como etiqueta de función principal,
- la tecla $+$ (tecla $(9,5)$),
- la tecla que tiene a 3 como etiqueta de función principal ($(9,4)$)
- y la tecla ENTER (tecla $(10,5)$).

Entonces, la calculadora desplegará el cálculo simbólico de esta expresión: $\sqrt{2}+3$; es decir, el cálculo exacto. Si se quiere saber una aproximación numérica de esta expresión, se puede hacer de tres maneras:



(a)

(b)

Figura 7-6 •

1. Después de haber capturado en la calculadora $\sqrt{2}+3$, como explicamos antes y oprimir la tecla ENTER, debemos oprimir la tecla de cambio derecho ($(9,1)$) y después la tecla ENTER, entonces la pantalla desplegará:

   $$\rightarrow\text{(NUM(ANS(1))}$$

   pues así hemos activado la función →NUM, que es la función contenida en la tecla ENTER resaltada con naranja en su calculadora, finalmente hay que oprimir otra vez la tecla ENTER ($(10,5)$) y la calculadora desplegará en su monitor la aproximación numérica de ese cálculo como se ilustra en la figura 7-6*(a)*.
2. Después de haber capturado en la calculadora $\sqrt{2}+3$, como explicamos antes, se debe oprimir la tecla (antes de oprimir ENTER) de cambio derecho $(9,1)$ y, posteriormente, la tecla ENTER, para activar su función →NUM y la calculadora desplegará la aproximación numérica como se ilustra en la figura 7-6*(b)*.
3. Capturar como hicimos antes $\sqrt{2.}+3.$ utilizando la tecla $(10,3)$ para el punto decimal. Si después de esto se oprime la tecla ENTER, la calculadora nuevamente desplegará la aproximación numérica solicitada.

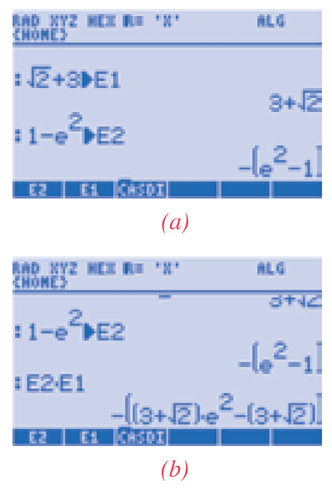
En general, cuando una expresión numérica contenga puntos decimales, la calculadora, después que se haya oprimido la tecla ENTER, dará como respuesta una aproximación decimal.

Almacenaje de objetos algebraicos

Para operar algebraicamente objetos, conviene almacenar éstos en la memoria de la calculadora y utilizarlos cuando sea conveniente. Supongamos que queremos almacenar en la memoria las expresiones $\sqrt{2}+3$ y $1-e^2$. Entonces escribimos, como explicamos antes, la primera expresión y se oprime la tecla F4 para activar la opción STO▶ del menú de herramientas; luego la pantalla de la calculadora mostrará la información

$$:\sqrt{2}+3\ \blacktriangleright$$

Esto significa que la calculadora espera la instrucción de qué etiqueta (nombre) se asignará a la expresión. Para ello se utiliza la tecla ALPHA ($(7,1)$), para escribir los caracteres alfanuméricos que se elegirán



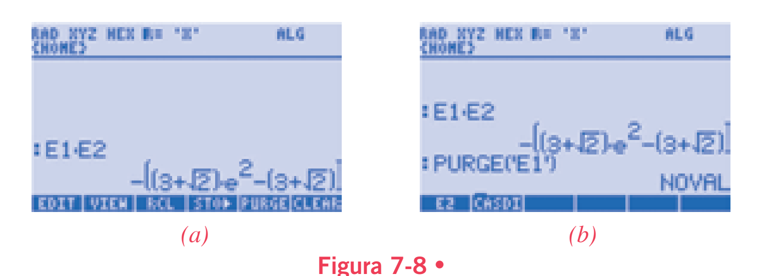
(a)

(b)

Figura 7-7 •

para el nombre de esta expresión; en nuestro caso se escogió E1 (mediante la cadena de activación de teclas $(7,1) \mapsto (1,5) \mapsto (9,2)$); después hay que presionar la tecla ENTER y el valor quedará almacenado en la memoria con el nombre E1. Se repite esta operación para almacenar la segunda expresión para la cual se eligió el nombre E2. Ahora, para utilizar estos objetos, se debe oprimir la tecla VAR, $(3,1)$. Entonces, el menú de herramientas de la calculadora cambiará a un menú donde aparecen ambos nombres de los objetos como se ilustra en la figura 7-7*(a)*. Para operar con ellos basta usar las teclas con las que se activan las opciones del menú de herramientas; por ejemplo, si oprimimos sucesivamente, para el caso de la pantalla mostrada en la figura 7-7*(a)*, las teclas F1, $\times$ y F2, la calculadora dará la respuesta mostrada en la figura 7-7*(b)*.

También es posible almacenar objetos algebraicos mediante la tecla STO ($(3,2)$) de manera completamente análoga. Por otra parte, la tecla TOOL, la tecla $(2,3)$, sirve para regresar al menú de herramientas previo. Como es natural, es posible eliminar de la memoria objetos que estén almacenados. Por ejemplo, si se quiere eliminar E1 que almacenamos



(a) *(b)*

Figura 7-8 •

previamente y la máquina se encuentra en el estado mostrado en la figura 7-7*(b)*, entonces se debe oprimir la tecla TOOL, para regresar al menú de herramientas mostrado en la figura 7-8*(a)*, después oprimir la tecla F5 para activar la opción PURGE, oprimir la tecla VAR ($(3,1)$), para regresar al menú de herramientas donde se encuentran los nombres de los objetos almacenados y oprimir, en este caso, la tecla F2 y después presionar la tecla ENTER para eliminar este objeto como se ilustra en la figura 7-8*(b)*.

7.1.5 Escritura de vectores y matrices en la Hp 50g

Existen dos formas de escribir vectores en la HP 50g. Una de ellas, que veremos más adelante, es un caso particular de la manera en la que se escriben las matrices con un editor especial para ese fin. La sintaxis de un vector $(a_1, a_2, \ldots, a_n)$ es la siguiente:

$$[\mathrm{a}_1, \mathrm{a}_2, \ldots, \mathrm{a}_\mathrm{n}]$$

donde [] es el par de paréntesis rectangulares que se activa oprimiendo primero la tecla de cambio izquierdo (tecla $(8,1)$) y después la tecla $\times$, (tecla $7,5$); los a_i son números u objetos algebraicos separados por comas; la coma se activa presionando la tecla cambio derecho, tecla $(9,1)$, y la tecla SPC, $(10,4)$. La figura 7-9*(a)* contiene el resultado en la pantalla de ejecutar la cadena de activación de teclas

para escribir el vector $(1,2,-5)$ en la HP 50g.

(a) *(b)*

Figura 7-9 •

La sintaxis para escribir una matriz es:

$$[[a_{11}, \ldots, a_{1n}], [a_{21}, \ldots, a_{2n}], \ldots, [a_{m1}, \ldots, a_{mn}]]$$

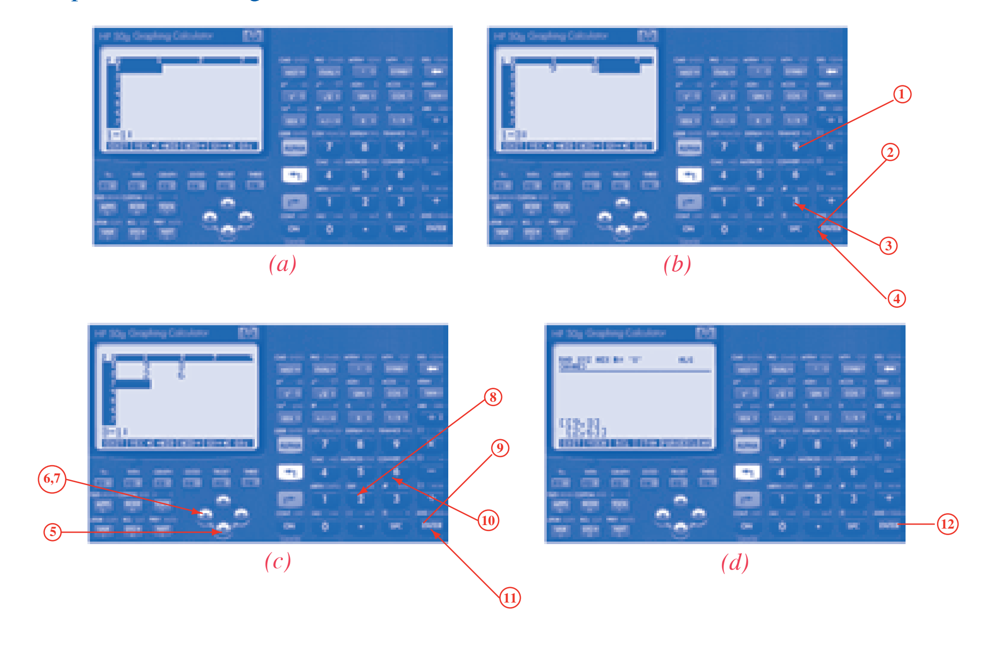
donde cada uno de los vectores $[a_{i1}, \ldots, a_{in}]$, $i = 1, \ldots, m$, son los vectores fila de la matriz. La figura 7-9*(b)* contiene en la pantalla el resultado de ejecutar la cadena de activación de teclas

para escribir la matriz $\begin{bmatrix} 1 & 2 \\ 1 & -3 \end{bmatrix}$.

Es importante que el lector tenga bien claro la sintaxis para escribir vectores y matrices en la calculadora. De esta manera $[1, 2, -5]$, se interpreta como una matriz columna, mientras que $[[1, 2, -5]]$ se traduce como una matriz fila.

Editor de matrices

La calculadora HP 50g cuenta con un editor especial de matrices. Para utilizar esta función hay que oprimir la tecla cambio izquierdo y después la tecla y así activar el menú que aparece en la figura 7-10*(a)*, entonces se oprime la tecla F5, si las entradas de la matriz se van a insertar por filas, o la tecla F6, si las entradas de la matriz se van a insertar por columnas. En el caso que ilustramos a continuación hemos elegido escribir por filas la matriz y, por tanto, las opciones quedan definidas como se ve en la pantalla de esa figura.



(a) *(b)* *(c)* *(d)*

Figura 7-10 •

Figura 7-11 •

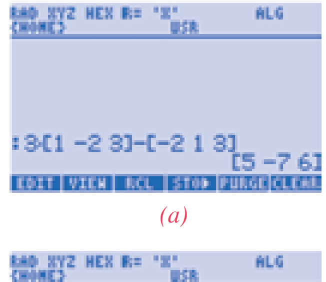
Este editor se utiliza como una hoja de cálculo electrónica. Por ejemplo, para escribir la matriz $\begin{bmatrix} 9 & 3 \\ 2 & 6 \end{bmatrix}$ primero oprimimos las teclas de acuerdo con el orden (1 a 4), que se indica en la figura 7-10*(b)*; con esto se introduce la primera fila y la pantalla mostrará la información que se ilustra en esa figura.[1] Para escribir la segunda fila se oprimen las teclas en el orden (5 a 11) al que se hace referencia en la figura 7-10*(c)*; al oprimir la última tecla la pantalla de la calculadora presentará la información de esa figura. Al presionar otra vez la tecla ENTER (paso 12 en la figura), la pantalla desplegará inmediatamente el formato que explicamos antes para escribir matrices y que se ve en la figura 7-10*(d)*. Finalmente, si se vuelve a oprimir la tecla ENTER, entonces la calculadora desplegará la matriz en modo algebraico como se presenta en la figura 7-11.

❍ **Nota 7.2** Para escribir un vector con el editor de matrices, se debe activar primero la opción VEC☐ del menú de herramientas del editor de matrices (cfr. figura 7-10*(a)*) y capturar el vector en la primera fila como se hizo antes.

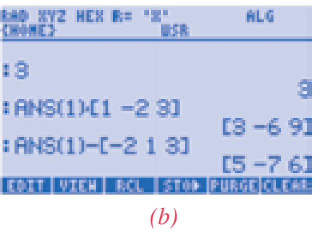
7.1.6 Operaciones con vectores

Antes de explicar el uso de la HP 50g para realizar algunas de las operaciones básicas entre vectores, invitamos al lector a que escriba y almacene en su calculadora los siguientes vectores, con las mismas etiquetas que utilizamos aquí, para usarlos más adelante:

$$\begin{aligned} u1 &= (1,-2,3) \\ u2 &= (-2,1,3) \\ u3 &= (1,-5,12) \\ u4 &= (0,-3,9) \\ v1 &= (-1,1,2,3,2) \\ v2 &= (1,0,-1,1,2) \\ v3 &= (1,-1,1,0,1) \end{aligned}$$



(a)



(b)

Figura 7-12 •

Suma, diferencia y multiplicación por un escalar

Para operar algebraicamente vectores con la HP 50g, simplemente se escriben en la calculadora y se emplean los signos usuales del álgebra para operar con éstos. Por ejemplo, para realizar la operación $3u1 - u2$, podemos escribir el número 3 (tecla (9,4)), oprimir la tecla para el producto (tecla (7,5)), escribir el vector $u1 = (1,-2,3)$, oprimir la tecla del signo menos (tecla (8,5)), escribir el vector $u2 = (-2,1,3)$ y, finalmente, oprimir la tecla ENTER para que la máquina despliegue el resultado como se ilustra en la figura 7-12*(a)*. Otra forma de hacerlo es escribir el número 3, oprimir la tecla ENTER, después la tecla ×, escribir el vector $u1$, oprimir ENTER, la tecla −, escribir el vector $u2$ y oprimir la tecla ENTER como se ilustra en la figura 7-12*(b)*.

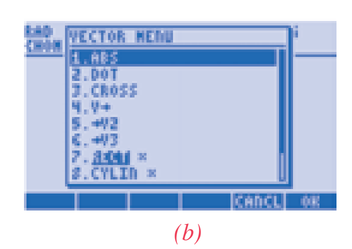
[1]Las imágenes de la HP 50g que se utilizan en esta figura no representan el modelo real de la calculadora; se han modificado sólo por comodidad gráfica.

Figura 7-13 •

Si se tienen almacenados los vectores $u1$ y $u2$, se puede realizar la misma operación oprimiendo la tecla VAR para activar el menú donde aparecen los objetos algebraicos que se han almacenado en la calculadora (cfr. el apartado de la pág. 765), entonces, se escribe el número 3, se oprimen la tecla ×, después la tecla que corresponda para activar el vector $u1$ de la memoria de la calculadora, en este caso F1, la tecla −, la tecla que corresponda para activar al vector $u2$ almacenado en la memoria, F2 y, por último, la tecla ENTER como se muestra en la figura 7-13.

Norma, producto punto y producto cruz (vectorial)

(a)

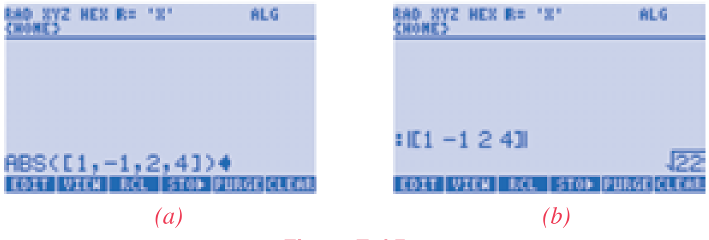


(b)

Figura 7-14 •

Utilizaremos un menú especial que tiene la HP 50g, el cual contiene estas funciones. Primero se tiene que activar el menú MTH oprimiendo la tecla cambio izquierdo y después la tecla $^{\text{MTH}}$SYMB (tecla (4, 4)), entonces aparece el menú que contiene la pantalla de la figura 7-14*(a)*, se selecciona con las teclas direccionales la opción VECTOR, se oprime la tecla F6 e inmediatamente la calculadora desplegará el menú que se muestra en la figura 7-14*(b)*. Las funciones ABS, DOT y CROSS son las que se emplean para calcular la norma (magnitud), el producto punto de dos vectores y el producto cruz (producto vectorial) de dos vectores (esta última función se aplica sólo a vectores con tres coordenadas), respectivamente.

Norma: Para calcular la norma de un vector, se selecciona la opción ABS con las teclas direccionales del menú mostrado en la figura 7-14*(b)*, se activa esta función con la tecla F6, se escribe el vector entre los paréntesis, como se ilustra en la figura 7-15*(a)* y se oprime la tecla ENTER, entonces la pantalla de la calculadora mostrará información como la que se ilustra en la figura 7-15*(b)*.



(a) *(b)*

Figura 7-15 •

Producto punto: Se activa la función DOT del menú VECTOR de la figura 7-14*(b)* y se escriben los vectores como se ve en la figura 7-16*(a)*; se oprime la tecla ENTER.

Producto cruz: Se activa la función CROSS del menú VECTOR mostrado en la figura 7-14*(b)* y se escriben los vectores como se ilustra en la figura 7-16*(b)*; se oprime la tecla ENTER.

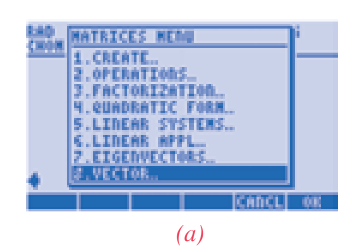
(a) *(b)*

Figura 7-16 •

❍ **Nota 7.3** 1. Toda función de la HP 50g tiene el formato FUN( ), por ejemplo la función ABS( ) mencionada antes, donde FUN es una función de algún menú o del teclado de la calculadora. Para

que esa función actúe sobre un argumento, entonces éste se debe escribir dentro de los paréntesis después del nombre que tiene dicha función; separados por comas si son más de un argumento (cfr. el uso de las funciones ABS, DOT y CROSS, explicadas arriba) y oprimir la tecla ENTER.

2. Si un objeto algebraico está almacenado en la memoria de la calculadora y se desea que cierta función actúe sobre él, se coloca el cursor entre los paréntesis de esa función y desde el menú VAR se activa la variable que contiene ese argumento.
3. Cualquier función de la HP 50g, ya sea de algún menú o del teclado, se puede utilizar escribiendo directamente su nombre en la pantalla de la calculadora y en seguida, entre paréntesis, escribir el argumento en el que se quiere evaluarla y después oprimir la tecla ENTER.



(a)

MATRIX VECTOR MENU
1.BASIS
2.CROSS
3.DOT
4.GRAMSCHMIDT
5.IBASIS
6.MATRICES..

(b)

Figura 7-17 •

Bases

La calculadora HP 50g cuenta con un menú que tiene funciones extras para actuar sobre conjuntos de vectores. Para utilizarlas hay que oprimir la tecla cambio izquierdo y la tecla MATRICES 5, tecla $(8,3)$, para activar el menú MATRICES que se muestra en la figura 7-17*(a)*; con las teclas direccionales seleccionar la opción VECTOR y oprimir la tecla F6 para activar este último menú, el cual se muestra en la figura 7-17*(b)*.

Las funciones DOT y CROSS ya fueron estudiadas antes. A continuación listamos las funciones y formas de uso para cada una de las opciones 1 y 5 de este menú:

BASIS: BASIS$(\{u,v,\ldots,w\})$ encuentra una base para el espacio generado por los vectores $u,v,\ldots,w$.

IBASIS: IBASIS$(\{u,v,\ldots,w\},\{t,\ldots,r\})$, halla una base para la intersección del espacio generado por los vectores $u,v,\ldots,w$ con el espacio generado por los vectores $t,\ldots,r$.

► **Ejemplo 7.1** Encontrar una base para el espacio generado por los vectores $u1$, $u2$, $u3$ y $u4$ que se almacenaron al inicio de este apartado en la memoria de la calculadora.◄

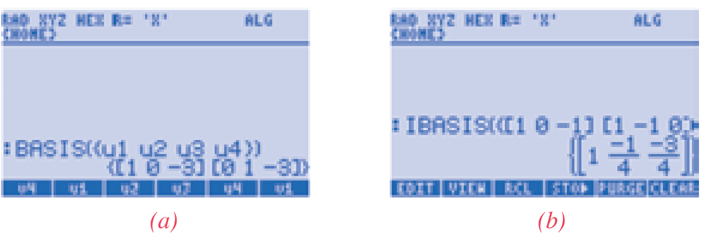
Solución Se escribe en la pantalla de la calculadora BASIS({u1,u2,u3,u4}), como se muestra en la figura 7-18*(a)* y la respuesta aparece ahí mismo. ✔

► **Ejemplo 7.2** Sean $S_1 = \text{gn}((1,0,-1),(1,-1,0))$ y $S_2 = \text{gn}((2,1,0),(-2,0,1))$. Hallar una base de $S_1 \cap S_2$.◄

Solución Se escribe en la pantalla de la calculadora

$$\text{IBASIS}(\{[1,0,-1],[1,-1,0]\},\{[2,1,0],[-2,0,1]\})$$

y se oprime la tecla ENTER como se ilustra en la figura 7-18*(b)*. La respuesta es, entonces, $\{(1,-1/4,-3/4)\}$. ✔



(a) *(b)*

Figura 7-18 •

7.1.7 Operaciones con matrices

Para ilustrar la forma en la que opera la calculadora HP 50g con matrices, guardaremos en la memoria de la máquina algunas matrices que utilizaremos después con los nombres que aquí les asignemos.

$$A=\begin{bmatrix}1 & -2 & 1\\ -2 & 1 & 3\end{bmatrix},\quad B=\begin{bmatrix}4 & 3 & -1\\ 2 & 5 & 3\end{bmatrix},\quad C=\begin{bmatrix}1 & 2 & -3\\ 2 & -1 & 2\\ 1 & 1 & 1\end{bmatrix},\quad D=\begin{bmatrix}1 & -2 & -3 & -1 & 1\\ -1 & -2 & 2 & 1 & 1\\ 2 & -1 & 1 & 4 & 2\\ 2 & 0 & -4 & -5 & 0\end{bmatrix}.$$

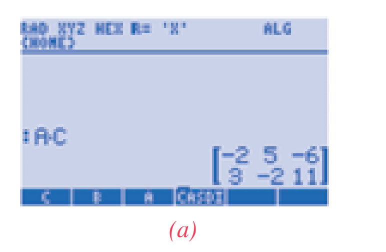
(*a*)

(*b*)

Figura 7-19 •

Suma, producto por un escalar y potencias de matrices

Estaremos usando constantemente el menú MATRICES, cuya activación se lleva a cabo con las teclas cambio izquierdo y la tecla $^{\text{MATRICES}}5$, tecla (8,3), cfr. figuras 7-17*(a)* y 7-4, y no indicaremos ya la manera de activar este menú ni los menús que contiene porque ya lo hemos explicado antes; tampoco haremos mención de cómo utilizar las funciones implementadas en la calculadora, pues ya tratamos antes ese tema (cfr. nota 7.3). Para operar algebraicamente matrices con la HP 50g se utilizan las teclas de producto (×), suma (+), resta (−) y de potenciación Y^{x}. Así, utilizando las matrices que almacenamos antes en la calculadora, tenemos que la operación $5A-3B$, se escribe en la calculadora como se ilustra en la figura 7-19*(a)* y al oprimir la tecla ENTER se obtiene el resultado que se muestra en la misma figura *(b)*. Mientras que las operaciones AC y C^3 se muestran en las figuras 7-20*(a)* y *(b)*, respectivamente.



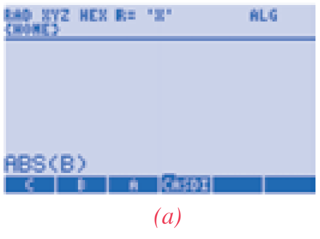
(*a*) (*b*)

Figura 7-20 •

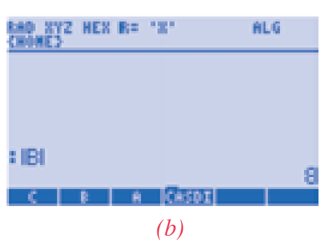
Norma

La función ABS del submenú OPERATIONS en el menú MATRICES (cfr. figura 7-17*(a)*, pág. 770), calcula la norma canónica de una matriz $M=[a_{ij}]\in\mathfrak{M}_{m\times n}$; así,

$$\|M\|=\left(\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}^{2}\right)^{1/2}.$$



(*a*)



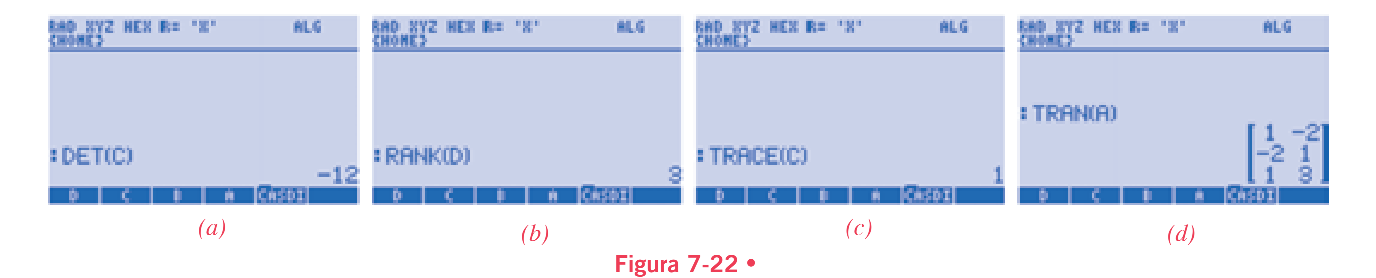
(*b*)

Figura 7-21 •

La figura 7-21*(a)* y *(b)* contiene la aplicación de la función ABS para el cálculo de la norma canónica de la matriz B que almacenamos en la memoria de la máquina. Otras normas se pueden encontrar en la opción NORMALIZE del submenú MATRIX del menú MTH; para tener acceso a éste último se tiene que ejecutar la cadena de activación de teclas [↰] [SYMB].

Determinante, rango, traza, transpuesta

La opción OPERATIONS del menú MATRICES (cfr. figura 7-17), contiene las funciones DET, RANK, TRACE y TRAN, éstas calculan, respectivamente, el determinante, el rango, la traza



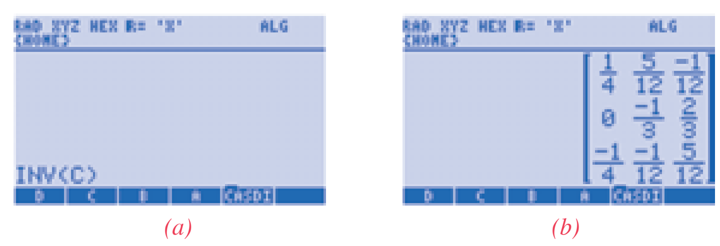
(a) *(b)* *(c)* *(d)*

Figura 7-22 •

y la transpuesta de una matriz. La figura 7-22*(a)*, *(b)*, *(c)* y *(d)* ilustra estas funciones aplicadas en algunas de las matrices que almacenamos en la memoria de la calculadora.

Inversa

La inversa de una matriz se calcula con la HP 50g mediante la función INV que se activa oprimiendo la tecla (tecla (6, 4)). En la figura 7-23*(a)* se ilustra el uso de esta función aplicado a la matriz C que se tiene almacenada en la memoria de la calculadora y el resultado que produce *(b)*.



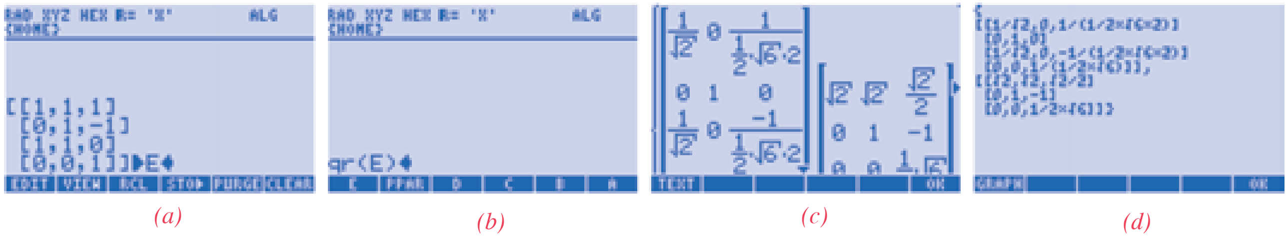
(a) *(b)*

Figura 7-23 •

7.1.8 *Factorización QR y ortogonalización, factorización LU*

Factorizaciones QR y LU

Para encontrar la factorización QR de una matriz E (cfr. apartado 4.1.4), se utiliza la función qr de la opción FACTORIZATION del menú MATRICES (cfr. figura 7-17). Por ejemplo, si se quiere encontrar la factorización QR de la matriz $E = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, se escribe esta matriz en la pantalla de la HP 50g; se almacena, digamos, con el mismo nombre; se activa la función qr; se escribe el nombre de la matriz, en nuestro caso con la tecla F1 del menú de herramientas VAR y se oprime la tecla ENTER; este proceso se encuentra plasmado en las figura 7-24*(a)*, *(b)*, *(c)* y *(d)*. En la última pantalla *(d)* se ha utilizado la



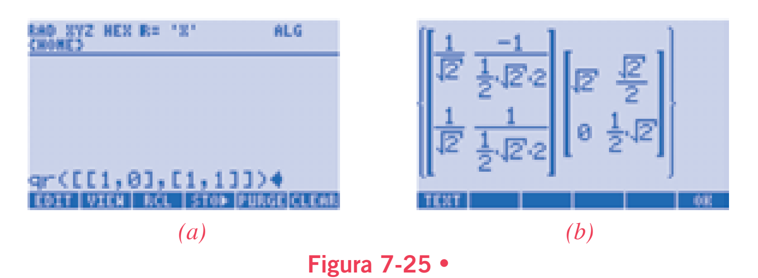
(a) *(b)* *(c)* *(d)*

Figura 7-24 •

opción TEXT, del menú de herramientas TOOL, para que se pueda visualizar el resultado completo en la pantalla en la sintaxis de la calculadora (cfr. ejemplo 4.33, para comprobar el resultado). De manera análoga se procede para la factorización LU, pero activando la función LU del mismo menú.

Ortogonalización

Para encontrar una base ortonormal de un subespacio generado por vectores linealmente independientes, basta encontrar la factorización QR, de la matriz que tiene como columnas a los vectores del conjunto generador, como se explicó en el párrafo anterior y las columnas de la matriz Q serán la base ortonormal correspondiente. Por ejemplo, la figura 7-25*(a)* y *(b)* contiene el proceso para encontrar una base ortonormal del espacio generado por el conjunto de vectores linealmente independientes $\{(1,1),(0,1)\}$.



(a) *(b)*

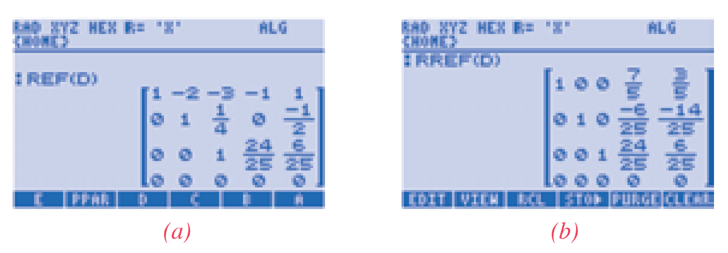
Figura 7-25 •

Por tanto, la base ortonormal está formada por las columnas de la matriz Q; es decir,

$$\{(1/\sqrt{2},1/\sqrt{2}),(-1/\sqrt{2},1/\sqrt{2})\}.$$

7.1.9 Forma escalonada y forma escalonada reducida con la HP 50g, sistemas lineales

El submenú LINEAR SYSTEMS del menú MATRICES (cfr. figura 7-17) contiene las funciones REF y RREF. La primera, cuando actúa sobre una matriz, da como resultado una forma escalonada; mientras que la segunda obtiene la forma escalonada reducida de dicha matriz. La figura 7-26*(a)* contiene el proceso de ejecución REF(D); mientras que la misma figura *(b)* contiene la ejecución RREF(D); donde D es la matriz que se tiene almacenada en la memoria de la calculadora.



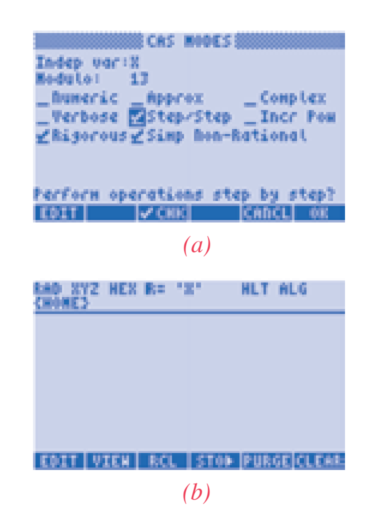
(a) *(b)*

Figura 7-26 •

Sistemas lineales

Para resolver sistemas lineales se pueden utilizar las funciones REF y RREF aplicadas a la matriz aumentada del sistema y aplicar, manualmente, sustitución regresiva.

7.1.10 Métodos de Gauss y Gauss-Jordan paso a paso en forma automática con la HP 50g



(a)

(b)

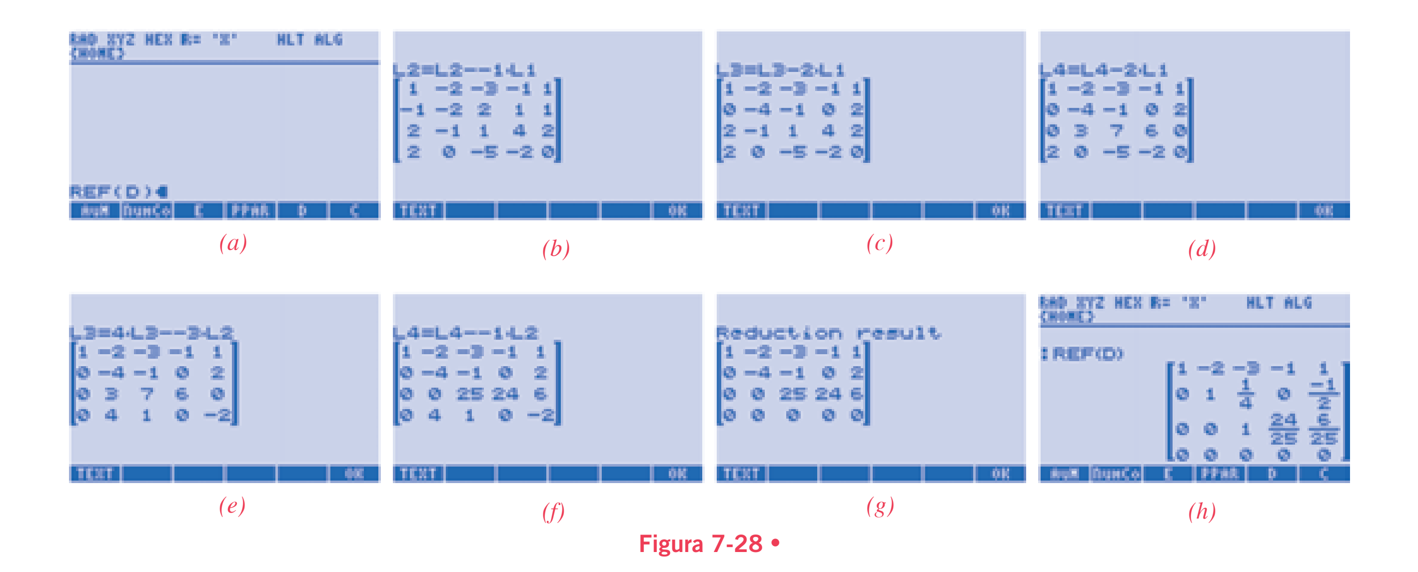
Figura 7-27 •

Activación del modo STEP/STEP

Con esta calculadora es posible efectuar las operaciones de renglón, paso a paso, y en forma automática para llevar una matriz a una forma escalonada y escalonada reducida. Para ello primero hay que oprimir la tecla MODE, después la tecla F3 para activar la opción CAS, con las teclas direccionales desplazar el cursor para selecionar la opción STEP/STEP, y activarla oprimiendo la tecla F3 para seleccionar la opción CHK, como se ilustra en la figura 7-27*(a)*; después hay que oprimir dos veces la tecla F6 y, entonces en la pantalla de la HP 50g se desplegará la leyenda HTL, junto a la leyenda ALG, indicando que dicha opción ya está activada como se muestra en la figura 7-27*(b)*.

Métodos de Gauss y Gauss-Jordan, paso a paso en forma automática

Una vez activado el modo STEP/STEP en la calculadora, sólo hay que aplicar las funciones REF y RREF del submenú LINEAR SYSTEMS del menú MATRICES, como se explicó en el apartado 7.1.9. Por ejemplo, si se desea encontrar una forma escalonada de la matriz D que se tiene almacenada en la calculadora, de tal manera que la máquina vaya desplegando en la pantalla cada una de las operaciones de renglón que efectúa al aplicar el método de Gauss, entonces se activa la función REF y se evalúa ésta en la matriz D; en la figura 7-28*(a)* se ilustra este proceso. Al oprimir la tecla ENTER, la pantalla de la calculadora desplegará la información contenida en la figura 7-28*(b)*. Observe que en la parte superior de la pantalla aparece el texto L2 = L2 −−1 · L1, que indica que al oprimir la tecla F6 para activar la opción OK del menú de herramientas, la calculadora efectuará la operación de renglón $R_2 \leftrightarrow R_2 - R_1$; el resultado de esa operación está contenido en la figura 7-28*(c)*. La figura 7-28 incisos *(c)* a *(g)* contiene cada uno de los pasos que la calculadora efectuará para obtener una forma escalonada de la matriz D; cualquiera de ellos se lleva a cabo oprimiendo la tecla F6 para activar la opción OK del menú de herramientas. Note que cuando la pantalla contiene la leyenda Reduction result se ha obtenido una matriz en forma escalonada equivalente a la matriz D *(g)*; si se vuelve a aceptar la opción OK oprimiendo la tecla F6, entonces la máquina realizará un último ajuste que consiste



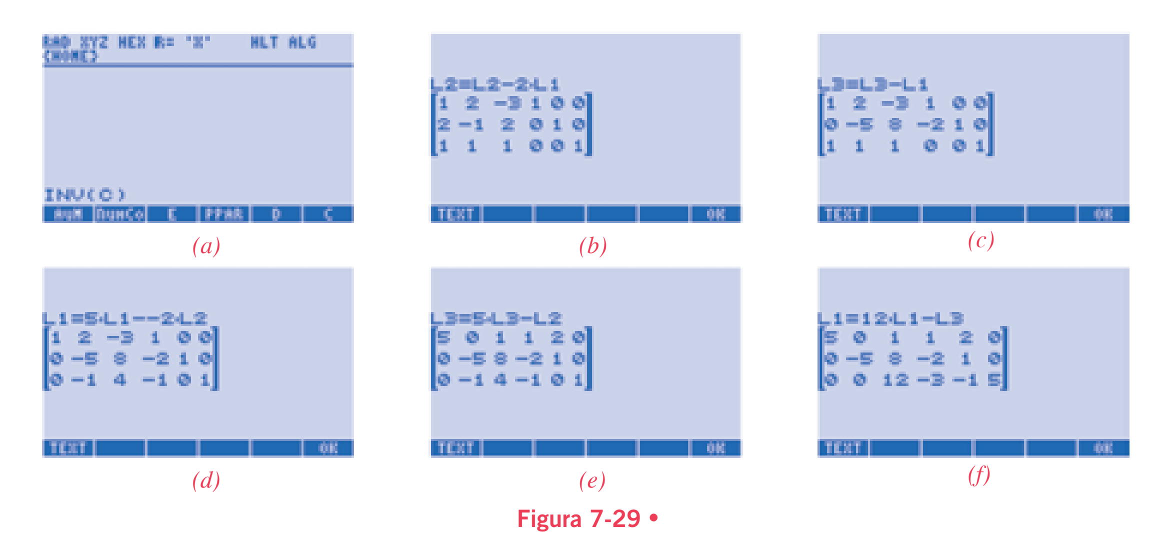
(a) (b) (c) (d) (e) (f) (g) (h)

Figura 7-28 •

en cambios de escala para que todos los pivotes sean unos, obteniendo el resultado mostrado en el inciso *(h)* de la misma figura (compare con el resultado ilustrado en la figura 7-26). Se hace lo mismo si se quiere obtener la forma escalonada reducida de una matriz, pero utilizando la función RREF en lugar de la función REF.

7.1.11 Inversa de una matriz paso a paso de manera automática con la calculadora HP 50g

De manera análoga a la descrita en el apartado anterior, se puede calcular la inversa de una matriz mostrando paso a paso el método de Gauss-Jordan que aplica la calculadora a la matriz ampliada $[C\ I_n]$, si aún se tiene la máquina en el modo STEP/STEP descrito antes. Para ello se utiliza la función INV (tecla $1/X$) aplicada a la matriz a la que se desea calcular su inversa. Las figuras 7-29 y 7-30 contienen este proceso utilizando la matriz C que tenemos almacenada en la memoria de la calculadora.

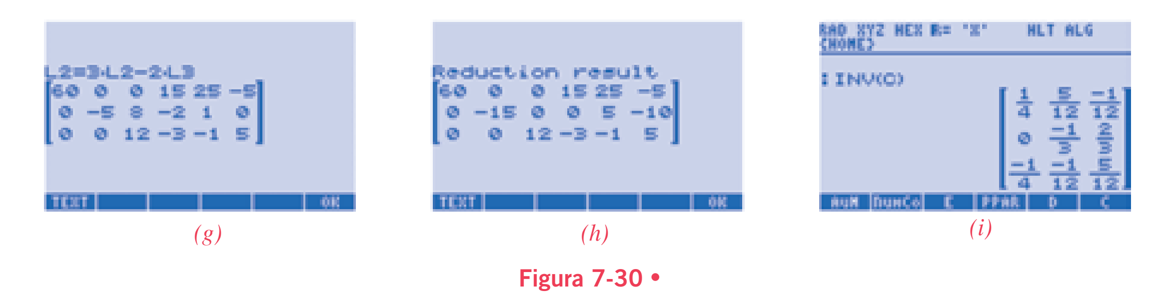


(a) *(b)* *(c)*

(d) *(e)* *(f)*

Figura 7-29 •

Como antes, la calculadora realiza el último ajuste, no de manera explícita, transformando en unos los elementos de la diagonal de la matriz en el lado izquierdo de la ampliación de la figura 7-30*(h)*, mediante cambios de escala y el resultado final se despliega en la pantalla de la figura 7-30*(i)*.

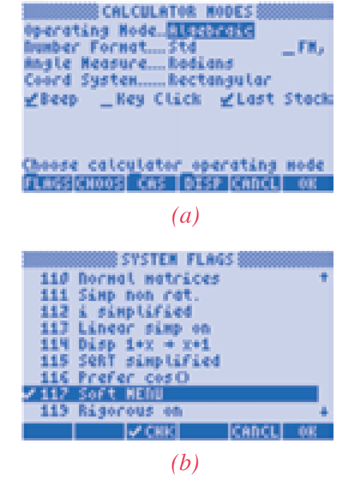


(g) *(h)* *(i)*

Figura 7-30 •

7.1.12 Métodos de Gauss y Gauss-Jordan con operaciones de renglón ejecutadas por el usuario

Es posible que el usuario dé instrucciones a la calculadora, por medio de algunos comandos, para ejecutar operaciones elementales de renglón a una matriz y encontrar una matriz equivalente escalonada

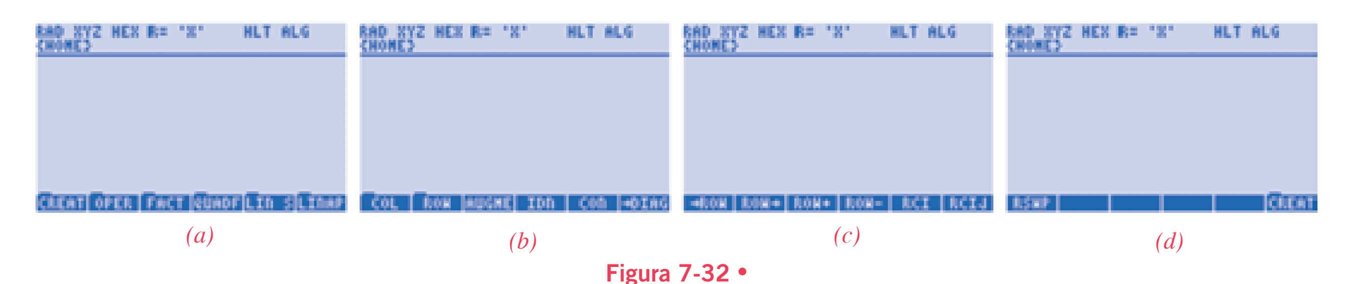


(a)

(b)

Figura 7-31 •

o la forma escalonada reducida de dicha matriz y, por tanto, poder calcular paso a paso la inversa de una matriz en forma manual. Para lograr este objetivo es conveniente tener la calculadora en un modo tal que las funciones aparezcan en el menú de herramientas en lugar de ser desplegadas en listas. Entonces, hay que oprimir la tecla MODE (tecla (2,2)), elegir la opción FLAGS oprimiendo la tecla F1, desplazarse en la pantalla con las teclas direccionales para selecionar la opción 117, Soft MENU, y oprimir la tecla F3 para activar ésta en el menú de herramientas como se indica en las figura 7.31*(a)* y *(b)*. Una vez que se haya oprimido dos veces la tecla OK, todos los menús que se elijan en la calculadora tendrán sus respectivos submenús y funciones como opciones en el menú de herramientas de la HP 50g. Necesitamos tres funciones que se encuentran en el menú MATRICES; por tanto, hay que oprimir la tecla cambio izquierdo, después la tecla MATRICES, la tecla F1 para activar el menú CREATE, figura 7-32*(a)* y la tecla F2 para activar el submenú ROW, figura 7-32*(b)*; inmediatamente el menú de herramientas lucirá como se muestra en la figura 7-32*(c)*. Las funciones que vamos a utilizar son RCI (F5), RCIJ (F6) y la tecla RSWP que se puede ver oprimiendo la tecla NXT, figura 7-32*(d)*. El efecto sobre



(a) (b) (c) (d)

Figura 7-32 •

(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j)

Figura 7-33 •

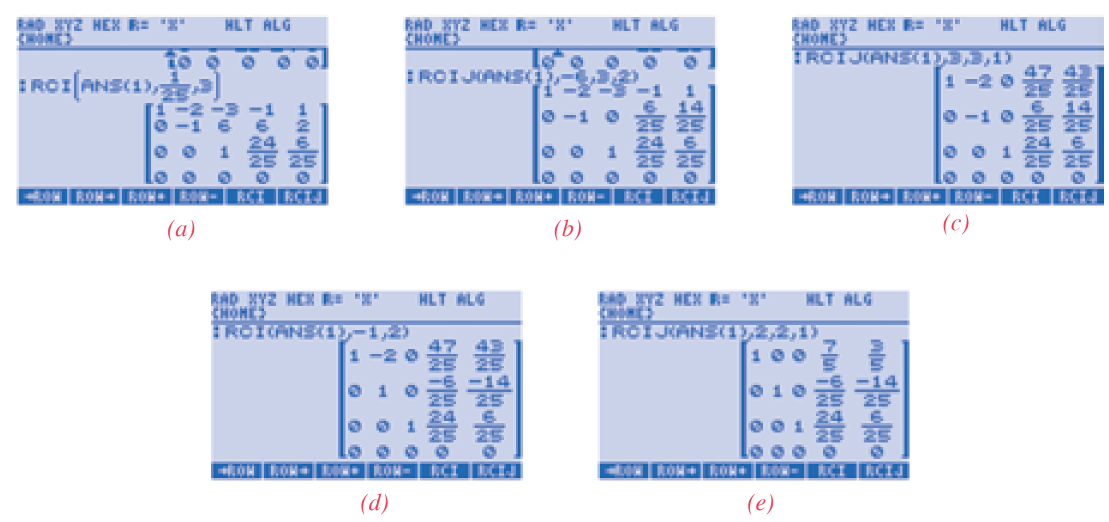


Figura 7-34 •

una matriz A y la sintaxis para producirlo al aplicar cada una de estas funciones en ésta es el siguiente:

- RSWP(A, I, J): intercambio de las filas I y J en la matriz A ($R_I \leftrightarrow R_J$).
- RcI(A, c, I): multiplicación de la fila I por el escalar c en la matriz A ($R_I \leftrightarrow cR_I$).
- RCIJ(A, c, I, J): cambiar la fila J de la matriz A por el resultado del producto del escalar c con la fila I de A y la suma de la fila J de A ($R_J \leftrightarrow cR_I + R_J$).

Como ejemplo, vamos a llevar a una forma escalonada la matriz D que se tiene almacenada en la memoria de la calculadora, que hemos hecho visible en la pantalla de la HP 50g, y dejamos las opciones del menú de herramientas como se muestra en la figura 7-33*(a)* (cfr. figura 7-32). En el primer paso haremos la operación de renglón $R_2 \leftrightarrow 1 \cdot R_1 + R_2$. Para ello oprimimos la tecla F6, activando así la función RCIJ, la tecla cambio izquierdo, la tecla ENTER (ANS) y se termina de escribir en la pantalla de la máquina el texto RCIJ(ANS(1),1,1,2), figura 7-33*(b)*; al oprimir la tecla ENTER la pantalla desplegará la matriz contenida en la figura 7-33*(c)*. La siguiente operación de renglón es $R_2 \leftrightarrow -2R_1 + R_2$; entonces oprimimos sucesivamente las teclas F6, cambio izquierdo, ENTER, para completar el texto mostrado en la figura 7-33*(d)*; después de oprimir la tecla ENTER, la calculadora desplegará la información de esta ejecución como se muestra en la figura 7-33*(e)*. Los siguientes pasos se ilustran en la figura 7-33*(f)* a *(j)*. Es evidente que se puede continuar de esta manera para aplicar el método de Gauss-Jordan y obtener la forma escalonada reducida de la matriz D. La figura 7-34*(a)*-*(e)* contiene los pasos necesarios para ello.

7.1.13 Inversa de una matriz por el método de Gauss-Jordan con operaciones de renglón ejecutadas por el usuario

Se puede calcular la inversa de una matriz C por el método de Gauss-Jordan con operaciones elegidas por el usuario, simplemente llevando a forma escalonada la matriz ampliada $[C\ I_n]$, donde I_n es la matriz identidad del mismo tamaño que la matriz C, siguiendo los pasos del apartado precedente. Puesto que es un poco fastidioso escribir, cada vez que se necesite, en la calculadora la matriz ampliada $[C\ I_n]$, vamos a programar un par de funciones en la HP 50g para que una vez que se haya escrito la matriz C, una

de ellas calcule la matriz ampliada. La primera, que llamaremos NumCol, es una función secundaria que sirve únicamente para indicarnos el número de columnas de una matriz sobre la que ella actúe y que utilizaremos en la programación de la función principal, AuM.

(a)

(b)

Figura 7-35 •

Programación de la función NumCol

Para programar la función NumCol, se escribe en la pantalla de la calculadora el guión que se muestra en la figura 7-35*(a)*. Primero se oprime la tecla cambio derecho (tecla $(9,1)$); después la tecla + , para activar los símbolos ≪≫; la tecla cambio derecho y la tecla 0, para activar la flecha →; la tecla ALPHA y la letra X; la tecla cambio izquierdo y la tecla ’, tecla $(4,3)$. La función $\to Q$ se encuentra en el menú CONVERT, en la opción REWRITE, que se activa oprimiendo la tecla cambio izquierdo, la tecla 6 ($(8,4)$), la tecla F4, la tecla NXT y la tecla F5 (o simplemente utilice las teclas cambio izquierdo y la tecla del número 0 para la flecha y escriba, sin espacio, la letra mayúscula Q); la función GET se encuentra en el menú MATRICES en la opción CREATE (o escriba directamente la palabra GET); la función SIZE se encuentra en el menú MATRICES en la opción OPERATIONS (o escriba directamente en la pantalla la palabra SIZE). Una vez que se ha editado el guión en la pantalla de la calculadora, se almacena en la memoria, utilizando la tecla STO, con el nombre NumCol. De esta manera, siempre que se quiera utilizar esta función se debe activar con la tecla VAR su nombre, inmediatamente escribir entre paréntesis la matriz sobre la cual se desea que actúe y oprimir la tecla ENTER para que ejecute. La función $\to Q$ convierte argumentos reales a números racionales simbólicos; la función SIZE da una lista con el tamaño de la matriz en la que se evalúa; la función GET(L,k) da el valor de la lista L que se encuentra en la posición k.

Programación de la función AuM

Una vez programada la función AuM, cuando se evalúe en una matriz C, esto es, cuando se oprima la tecla ENTER después de AuM(C), se imprimirá en la pantalla la matriz aumentada $[C\ I_n]$. De manera completamente análoga a como se procedió con la función anterior, se escribe el guión que se muestra en la figura 7-35*(b)*. En este caso la función TRAN se encuentra en el menú MATRICES en la opción OPERATIONS, que ya hemos descrito antes, y actúa en una matriz calculando su transpuesta; la función IDN(k) da como resultado la matriz identidad de orden k y se encuentra en el menú MATRICES en su opción CREATE; la función AUGMENT(A,L) adjunta al arreglo A el arreglo L, que se encuentra en el mismo menú que la función IDN. Con la tecla STO se almacena en la memoria con el nombre AuM.

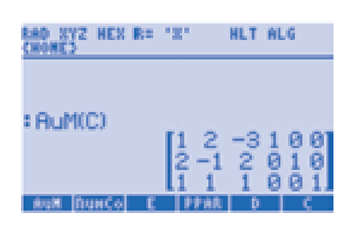


Figura 7-36 •

Inversa de una matriz con operaciones de renglón seleccionadas por el usuario

Vamos a calcular paso a paso la inversa de la matriz C que se tiene registrada en la memoria de la calculadora. Para ello, primero creamos la matriz aumentada (ampliada) $[C\ I_3]$ con la función AuM, como se ilustra en la figura 7-36; después aplicamos el método de Gauss-Jordan con operaciones de renglon elegidas por el usuario, descrito en el apartado anterior, hasta obtener la matriz $[I_n\ C^{-1}]$. Los pasos de este proceso se muestran en la figura 7-37*(a)*-*(i)* (recuerde tener activado el menú MATRICES-CREATE-ROW en el menú de herramientas con la opción Soft MENU); y la inversa de la función se encuentra en la misma figura *(j)*.

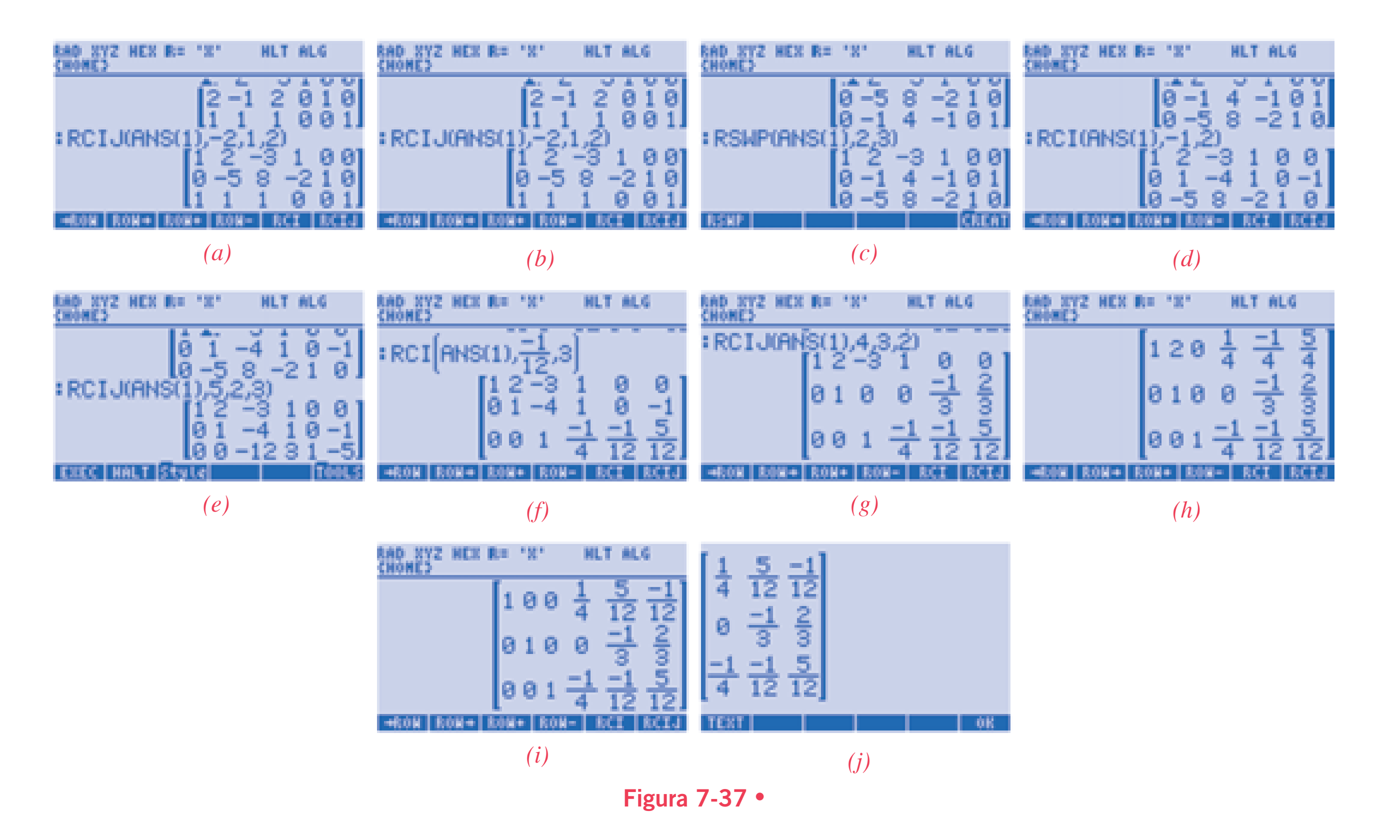


Figura 7-37 •

7.1.14 Transformaciones lineales, núcleo e imagen

Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal con representación matricial $A \in \mathfrak{M}_{m\times n}$ respecto a las bases canónicas. Con la calculadora HP 50g es posible encontrar bases para la imagen, $T(\mathbb{R}^n)$, y para el núcleo, $\ker(T)$, de la transformación lineal; que corresponden, respectivamente, al espacio columna y al espacio nulo de la matriz A. Las funciones de la calculadora que se utilizan para este fin son IMAGE y KER; ambas se localizan en el submenú LINEAR APPL del menú MATRICES. La figura 7-38 *(a)* contiene el proceso de aplicar la función IMAGE a la matriz D que tenemos almacenada en la memoria de la calculadora; mientras que la misma figura, *(b)* y *(c)*, contiene el proceso de aplicar a esta matriz la función KER; se tienen que utilizar las flechas direccionales para poder ver el resto de la información en la pantalla (misma figura *(d)*).

Figura 7-38 •

7.1.15 Valores y vectores propios

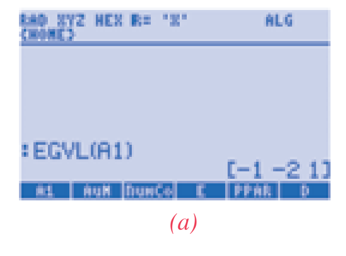
El submenú EIGENVECTORS del menú MATRICES contiene funciones para calcular el polinomio característico, valores y vectores propios de una matriz. A continuación se describen la sintaxis y efecto que produce cada una de esas funciones.

- EGVL(A) calcula los valores propios de la matriz A.
- EGV(A) produce un par de matrices: la primera contiene los vectores propios de la matriz A como columnas; la segunda contiene los valores propios de la matriz A, que corresponden a los vectores propios columna, de la primera matriz.
- PCAR(A) calcula el polinomio característico de la matriz A.

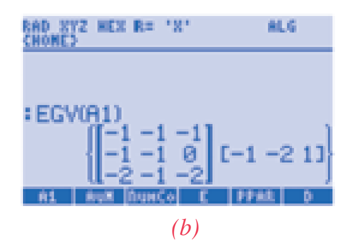
Por ejemplo, si escribimos y guardamos en la memoria de la calculadora la matriz

$$A1 = \begin{bmatrix} -1 & -2 & 1 \\ -2 & 1 & 1 \\ 2 & -4 & 0 \end{bmatrix}$$

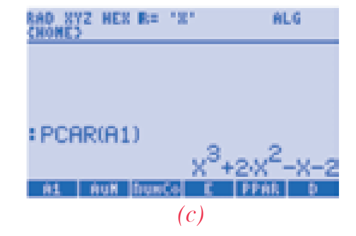
con el mismo nombre, la figura 7-39 *(a)*, *(b)* y *(c)* contiene la aplicación de cada una de estas funciones a la matriz $A1$.



(a)



(b)



(c)

Figura 7-39 •

❍ **Nota 7.4** Es posible que la calculadora dé un conjunto vacío de valores o vectores propios para cierto tipo de matrices, si se está trabajando en el modo simbólico; de suceder así, se debe utilizar las teclas de cambio derecho y la tecla ENTER para que la máquina despliegue el resultado en modo numérico o poner la calculadora en modo Approx. En general, este proceso se debe llevar a cabo siempre que la HP 50g dé una respuesta vacía si está trabajando en modo simbólico (algebraico).

7.1.16 Números complejos con la HP 50g

Para trabajar con números complejos en la calculadora HP 50g, únicamente se debe permitir que la máquina opere con este tipo de números. Para ello se debe oprimir la tecla MODE, después la tecla F3 para seleccionar la opción CAS, seleccionar con las teclas direccionales la opción _Complex y la tecla F3 para aceptar esta opción y, finalmente, la tecla F6 (OK) (cfr. figura 7-5*(d)*, pág. 764). Una vez que se ha activado el modo Complex, la calculadora aceptará y desplegará información con números complejos. Para escribir un número complejo en la calculadora se emplea la sintáxis $a + bi$ o (a, b) y todas las operaciones algebraicas se llevan a efecto de manera completamente análoga a como se hizo con números reales.

7.2 MATLAB y álgebra lineal

MATLAB es un potente paquete computacional interactivo que trabaja sobre arreglos matriciales directamente; de hecho sus siglas significan *Matrix Laboratory*. Ésta es la principal ventaja que tiene este sistema, pues le permite resolver muchos problemas técnicos computacionales en una fracción del tiempo que se necesitaría para escribir un programa en un lenguaje como C o FORTRAN. Las capacidades de este paquete van más allá de trabajar únicamente con arreglos matriciales. MATLAB es un lenguaje de alto rendimiento que integra computación, visualización y programación en un ambiente *amigable* que utiliza la notación matemática como base. Entre las muchas aplicaciones que tiene MATLAB están el desarrollo de algoritmos, modelación, simulación, análisis de datos, gráficas, matemáticas y computación. Para resolver necesidades específicas en diversas áreas, MATLAB ha desarrollado las llamadas *toolboxes*; que son conjuntos de herramientas diseñados convenientemente para satisfacer esas necesidades. Las versiones para estudiantes por lo general incluyen tres *toolboxes*: *Signal Processing Toolbox*, *Control System Toolbox* y *Symbolic Math Toolbox*. Esta última es una caja de herramientas que permite realizar matemáticas en forma simbólica. El espacio que ocupa en la memoria de una computadora es relativamente pequeño en comparación con otros paquetes y su precio comercial es más económico. La programación en MATLAB es muy sencilla y una enorme ventaja es la interface gráfica que con ella puede utilizarse. Estas bondades y las cajas de herramientas han hecho de MATLAB un paquete de gran uso en ingeniería y ciencias. Además de las versiones para estudiantes, también existen las versiones profesionales que, como es natural, son más poderosas (y más caras); en este texto utilizaremos una versión para estudiantes y trataremos únicamente lo que concierne a la parte de álgebra lineal en un contexto muy básico.

7.2.1 Interacción con MATLAB y almacenamiento de información

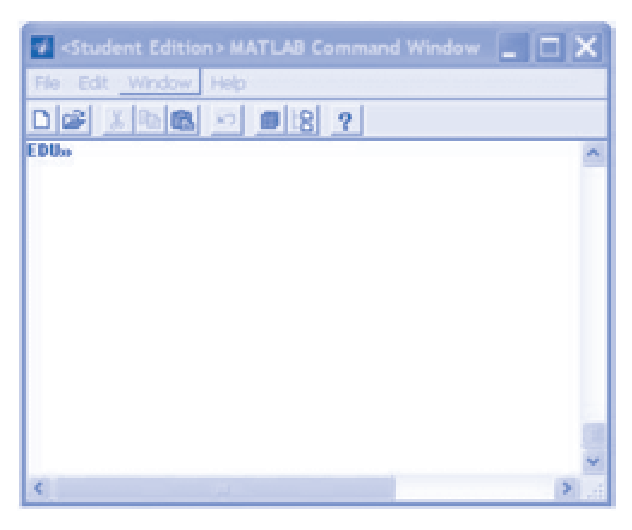


Figura 7-40 •

Una vez que se ha instalado una de las versiones del paquete en la computadora y se ha activado el mismo, la computadora desplegará (si el sistema operativo trabaja sobre Windows) una ventana como la que se muestra en la figura 7.40 (<Student Edition> MATLAB Command Window). Éste es el medio con el cual el usuario interactúa con MATLAB; las instrucciones que a través del teclado se escriban aparecerán después de EDU» en esa ventana. De esta manera, si se escribe con el teclado el número 3 y se oprime la tecla ENTER (INTRO) de la computadora:

```
EDU» 3
ans =
     3
EDU»
```

la computadora habrá almacenado el valor 3 con la etiqueta (nombre) `ans`; que es el nombre que por defecto asigna MATLAB a ese objeto. Esto significa que si en lugar de ello se hace

```
EDU» a=5
a =
     5
EDU»
```

en la memoria de la calculadora habrá quedado almacenado el valor 5 con el nombre a. Este valor quedará todo el tiempo con esa etiqueta asignada hasta que el usuario le reasigne un nuevo valor a dicha etiqueta. Pero, si se desactiva MATLAB (escribiendo en la ventana la palabra exit y oprimiendo después la tecla ENTER, o simplemente cerrando la ventana de comandos de MATLAB), entonces todos los objetos almacenados quedarán borrados de la memoria y, cuando nuevamente el usuario utilice MATLAB, ya no podrá hacer uso de ellos. Sin embargo, con la instrucción save, se pueden almacenar permanentemente los objetos en la memoria de la computadora y, con la instrucción load, podrá cargarlos otra vez para utilizarlos. La sintaxis para cada uno de estos comandos es:

- save *nombre del objeto*
- load *nombre del objeto*

Supongamos, por ejemplo, que vamos a cerrar MATLAB y deseamos almacenar el objeto a; escribimos en la ventana de MATLAB y después oprimimos la tecla ENTER de la computadora:

```
EDU» save a
EDU»
```

Si salimos del paquete y después escribimos las instrucciones

```
EDU» load a
EDU»
```

la variable a volverá a cargarse para utilizarla nuevamente.

❍ **Nota 7.5** 1. MATLAB ejecuta una instrucción sólo si después de escribirla en su ventana se oprime la tecla ENTER. Ya no haremos explícito este detalle de aquí en adelante.

2. Está permitido usar cualquier carácter alfanumérico, incluso el guión bajo, para nombrar objetos en MATLAB (sin dejar espacios en blanco).

3. La sintaxis de MATLAB es sensible a mayúsculas y minúsculas; por ejemplo, las etiquetas meNor y menor son distintas en este programa.

Podemos saber las variables (objetos) que tenemos almacenadas en cada sesión utilizando el comando who:

```
EDU» who
Your variables are:
a
EDU»
```

y ver las variables que tenemos permanentemente almacenadas con el comando what:

```
EDU» what
MAT-files in the current directory C:\MATLAB\BIN\libro
a
EDU»
```

El paquete indica el directorio donde está almacenada la variable `a` —directorio en curso donde se debe estar trabajando para poder tener acceso a esa variable, en nuestro caso `C:\MATLAB\BIN\libro`—. De no ser así, se debe cambiar de directorio desde la ventana de MATLAB (consulte la guía del usuario). Las instrucciones que se den a MATLAB en una sesión no se pueden almacenar, únicamente se pueden guardar objetos numéricos y algebraicos. Para ver las variables que se encuentran en el directorio en curso se escribe el comando `dir`. Para borrar una variable se utiliza la instrucción `delete` seguida (con un espacio) del nombre del archivo; así, `delete a.mat` borraría de la memoria el objeto `a`. Existen formas de almacenar ya sea en forma de texto, únicamente para registrar por escrito, o en forma de guión, para ejecutar y modificar cuantas veces sea necesario un conjunto de instrucciones. La primera es la instrucción con sintaxis `diary` *nombre de archivo* con extensión `txt` y la segunda es utilizando un archivo `m` que veremos más adelante.

Uso del punto y coma para evitar despliegue de información

Para cerrar este apartado es conveniente señalar una característica muy útil en la sintaxis de MATLAB. Cuando se escribe un punto y coma después de una instrucción, el programa la ejecuta, pero no la despliega; éste es evidentemente un rasgo muy útil de MATLAB cuando no se desea visualizar en la pantalla información innecesaria.

```
EDU» valor_de_X=3.1618;
EDU»
```

7.2.2 Escritura de matrices y operaciones básicas

Para escribir una matriz, se encierran entre paréntesis rectangulares las componentes de la matriz separadas por un espacio cada una (o por una coma) y las filas en distintas líneas o separadas por el carácter punto y coma

```
EDU» A=[ 1 -2 3
2 1 1
-1 1 0]
A =
     1    -2     3
     2     1     1
    -1     1     0
EDU»
EDU» A=[1,2,3;2,1,1;-1,1,1]
A =
     1     2     3
     2     1     1
    -1     1     1
EDU»
```

Las operaciones básicas en MATLAB utilizan la notación usual de una calculadora:

- + para la suma `a+b`.
- - para la diferencia `a-b`.
- * para el producto `a*b`.
- / para la división `a/b`.
- sqrt para la raíz cuadrada `sqrt(a)` $= \sqrt{a}$.
- ^ para elevar a una potencia `a^b` $= a^b$.
- \ para la división `a\b` $= b/a$.

```
EDU» -3+5
ans =
      2
EDU» 4*sqrt(2)
ans =
    5.6569
EDU» 4/2
ans =
      2
EDU» 4\2
ans =
    0.5000
EDU» 2^(1.8)
ans =
    3.4822
EDU» A=[1 -1 2; -2 1 3]; B=[0 1 -1; 1 2 1];
EDU» C=[-1 1 3; 2 2 1; -3 2 1];
EDU» 3*A-2*B
ans =
      3     -5      8
     -8     -1      7
EDU» A*C
ans =
     -9      3      4
     -5      6     -2
EDU» C^2
ans =
     -6      7      1
     -1      8      9
      4      3     -6
EDU» D=3*C^3;E=(A-2*B)*D
E =
    381  -291   -48
    -54  -423    21
EDU»
```

Una característica muy importante de MATLAB es que no se requiere declarar el tipo de variables con las que se trabaja; así, se puede hacer cálculos con números reales y complejos en el mismo ambiente sin hacer ningún cambio o declaración inicial.

```
EDU» (3^2-25)^(1/2)
ans =
   0.0000+ 4.0000i
EDU»
```

7.2.3 Formatos y modo simbólico

MATLAB tiene distintos formatos para desplegar información numérica. Algunos son:

- `format short e`: cinco dígitos más exponente; por ejemplo, `3.5833+e01`.
- `format long`: 16 dígitos.
- `format long e`: 16 dígitos más exponente.
- `format bank`: 2 dígitos decimales; por ejemplo `12.45`.
- `format rat`: aproximación racional; por ejemplo `13/125`.
- `format` o `format short`: el formato que por defecto tiene MATLAB y que es con el que se hicieron los cálculos del apartado precedente.

```
EDU» 1/3
ans =
    0.3333
EDU» format short e
EDU» 1/3
ans =
  3.3333e-001
EDU» format rat
EDU» 4/128
ans =
     1/32
EDU» format long
EDU» 1/3
ans =
   0.33333333333333
EDU» format
EDU» 4/128
ans =
    0.0312
EDU»
```

Las versiones para estudiantes de MATLAB contienen el *toolbox* de matemática simbólica. Para hacer operaciones en modo simbólico, requerimos declarar las literales que se van a utilizar con el comando `syms`, separando cada literal con un espacio en blanco:

```
EDU» syms x
EDU» J=[x x^2
-x 2*x]
```

```
J =
[    x, x^2]
[   -x, 2*x]
EDU» J^2
ans =
[     x^2-x^3,       3*x^3]
[      -3*x^2, -x^3+4*x^2]
EDU»
```

El comando `sym` transforma una expresión al modo simbólico y todo lo que se opere con ella estará en modo simbólico. Para ilustrarlo, emplearemos la matriz `A` que tenemos almacenada en la memoria.

```
EDU» S=sym(A)
S =
[  1, -1,  2]
[ -2,  1,  3]
EDU» sqrt(2)*S
ans =
[     2^(1/2),     -2^(1/2),   2*2^(1/2)]
[ -2*2^(1/2),       2^(1/2),   3*2^(1/2)]
EDU»
```

En ocasiones es difícil leer las expresiones en este modo, se puede utilizar un comando que ayuda un poco a visualizar mejor la información desplegada; el comando `pretty`:

```
EDU» pretty(ans)
                         [   1/2         1/2         1/2]
                         [ 2           -2          2 2    ]
                         [                                ]
                         [     1/2      1/2          1/2]
                         [-2  2        2           3  2    ]
EDU»
```

7.2.4 Matrices especiales, información básica y edición de matrices

Matrices especiales

Para no tener que escribir una matriz como la matriz cero o la matriz identidad, MATLAB tiene algunos comandos diseñados para este propósito, algunos son:

- Matriz cero de orden $m \times n$: `zeros(m,n)`.
- Matriz identidad de orden n: `eyes(n)`.
- Matriz de unos: `ones(m,n)`

```
EDU» zeros(2,3)
ans =
     0     0     0
     0     0     0
EDU» eyes(3)
ans =
     1     0     0
     0     1     0
     0     0     1
EDU» ones(3,2)
ans =
     1     1
     1     1
     1     1
EDU»
```

Información básica y edición de matrices

El comando `size` sirve para conocer el tamaño de una matriz; mientras que el comando `length` se utiliza para conocer el número de componentes de un vector.

```
EDU» M=[1 -1 2; 2 1 1]; u=[1;2;3;4]; v=[1,-1,2,0,1];
EDU» size(M)
ans =
     2     3
EDU» size(u)
ans =
     4     1
EDU» size(v)
ans =
     1     5
EDU» length(u)
ans =
     4
EDU» length(v)
ans =
     5
EDU»
```

Si se ha etiquetado una matriz con el nombre `M`, entonces se tienen los siguientes comandos para visualizar componentes, filas y columnas de esta matriz:

- `M(i,j)`: despliega la información que contiene la componente que se encuentra en la fila `i` y en la columna `j` de la matriz `M`.
- `M(i,:)` despliega la fila `i` de la matriz `M`.
- `M(:,j)` despliega la columna `j` de la matriz `M`.

```
EDU» M=[-1 2 3 4
-1 1 2 3
0 1 1 1]

M =
    -1     2     3     4
    -1     1     2     3
     0     1     1     1
EDU» M(2,3)
ans =
     2
EDU» M(3,:)
ans =
     0     1     1     1
EDU» M(:,4)
ans =
     4
     3
     1
EDU»
```

Por supuesto, se puede editar una entrada o una columna o una fila de la matriz `M` con estas instrucciones.

```
EDU» M(1,2)=-5
M =
    -1    -5     3     4
    -1     1     2     3
     0     1     1     1
EDU»  M(2,:)=[-1 0 1 1]
M =
    -1    -5     3     4
    -1     0     1     1
     0     1     1     1
EDU»
```

Un detalle que puede resultar útil en la edición de matrices, es que en MATLAB es muy simple adjuntar una matriz a otra para ampliar la primera. Utilizando la matriz `M` que ya tenemos escrita, vamos a ampliarla con la matriz `M2` que tenemos que escribir en la ventana de MATLAB.

```
EDU» M2=[ -3 1
2 2
0 1]
M2 =
    -3     1
     2     2
     0     1
```

```
EDU» K=[M M2]

K =
    -1    -5     3     4    -3     1
    -1     1     2     3     2     2
     0     1     1     1     0     1
```

7.2.5 Operaciones de renglón con MATLAB

Con los comandos descritos antes se puede llevar a cabo operaciones de renglón para encontrar, por ejemplo, una forma escalonada equivalente.

```
EDU» M(1,:)=-1*M(1,:)

M =
     1     5    -3    -4
    -1     0     1     1
     0     1     1     1
EDU» M(2,:)=M(1,:)+M(2,:)

M =
     1     5    -3    -4
     0     5    -2    -3
     0     1     1     1
EDU» M(3,:)=M(2,:)-5*M(3,:)

M =
     1     5    -3    -4
     0     5    -2    -3
     0     0    -7    -8
EDU»
```

El intercambio de filas es un poco más laborioso:

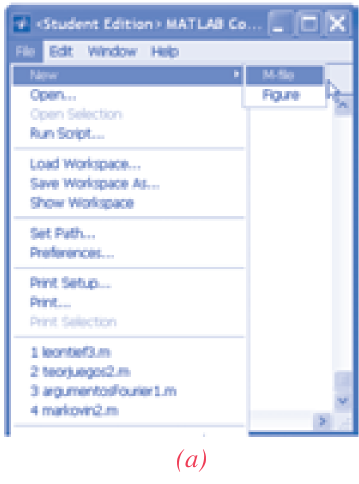
```
EDU» M1=[1 5 -3 -4
0 5 -2 -3
0 1 1 1]

M1 =
     1     5    -3    -4
     0     5    -2    -3
     0     1     1     1
EDU» I=M1(2,:);J=M1(3,:);
EDU» M1(2,:)=J; M1(3,:)=I

M1 =
     1     5    -3    -4
     0     1     1     1
     0     5    -2    -3
EDU»
```

Más adelante veremos cómo es posible hacer las operaciones de renglón en forma más simple programando una función en MATLAB.

❍ **Nota 7.6** Con las teclas direccionales de la computadora es posible volver a colocar información en la pantalla de MATLAB sin tener que teclearla; con la de dirección hacia arriba se obtiene la información precedente en forma sucesiva y, con la tecla de dirección hacia abajo, la información posterior. Este detalle resulta muy útil para repetir y editar instrucciones sin tener que escribirlas por completo.

7.2.6 Funciones programadas por el usuario, programación en MATLAB y operaciones de renglón

En MATLAB es posible que el usuario programe sus propias funciones. Para ello hay que editar el guión del programa que define la función en un editor y guardarlo, con extensión `m`, en el directorio donde se vaya a utilizar. MATLAB cuenta con un editor especial para este fin; para activarlo, hay que hacer clic en la opción FILE de la ventana de MATLAB, después en New y M-file como se ilustra en la figura 7-41*(a)*; entonces se abrirá una nueva ventana con el editor de MATLAB como se muestra en la figura 7-41*(b)*.



(a)

(b)

Figura 7-41 •

En la figura 7-42 está escrito un programa para calcular el polinomio característico de una matriz de tamaño 2×2 en el editor de texto que acompaña a MATLAB. A continuación explicamos a grandes rasgos las partes de las que está formado este programa.

1. La palabra reservada `function`, instrucción que se utiliza para indicar a MATLAB que el programa es para definir una función.
2. Una etiqueta de salida, en este caso utilizamos la letra `p`, que sigue de la palabra reservada `function`.
3. Después del signo igual, que sigue de la etiqueta de salida (para este ejemplo `p`), viene el nombre de la función. Hemos llamado a esta función `policarac`.
4. Inmediatamente después del nombre que se usa para definir la función, entre paréntesis se escriben las variables de las cuales va a depender la función, separadas por comas; en este ejemplo la única variable es `A`.

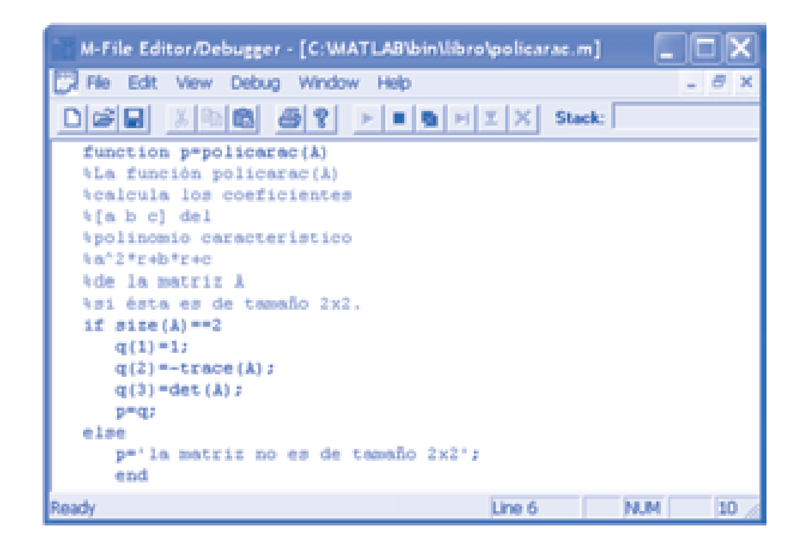


Figura 7-42 •

5. El comando de decisión lógica `if` y partes que lo conforman: `else` y `end`.
6. El comando de comparación lógica `==`. El comando `if` indica al programa que ejecute las instrucciones que están debajo de él y antes de `else` si los miembros izquierdo y derecho de `==` son iguales; en caso contrario, el comando `if` indica al programa que ejecute las instrucciones que están por debajo de `else`; en este caso desplegar el texto entre comillas.
7. La instrucción `q(1)=1` define la primera componente del vector `q` igual a 1.
8. La instrucción `q(2)=-trace(A)` define la segunda componente del vector `q` igual a la evaluación de la función `trace` en la variable `A`. `trace` es una función de MATLAB que calcula la traza de la matriz en la que se evalúa.
9. La instrucción `q(3)=det(A)`, define la tercera componente del vector `q` igual a la función `det` evaluada en la variable `A`. `det` es una función de MATLAB que calcula el determinante de la matriz en la que se evalúa.
10. La asignación del valor `q` a la variable `p` mediante la instrucción `p=q`.
11. Todos los textos después del signo `%` son comentarios que no ejecuta el programa. Sin embargo, todo lo que se escriba después de ellos, al inicio del programa, se desplegará en la ventana de MATLAB si se ejecuta la instrucción `help policarac`; por tanto, estos comentarios sirven para ayudar al programador a describir la función y su uso.

Por último, el archivo debe guardarse con el mismo nombre que tiene la función con la extensión, asignada por defecto en este editor, `m`; en este caso `policarac.m`. Veamos cómo trabaja la función `policarac.m`:

```
EDU» help policarac
La función policarac(A)
calcula los coeficientes
[a b c] del
polinomio característico
a^2*r+b*r+c
de la matriz A
si ésta es de tamaño 2x2.
```

```
EDU» C=[-1 3
1 2]
C =
    -1     3
     1     2
EDU» policarac(C)
ans =
     1    -1    -5
EDU» D=[1 1;-1 2; 3 1];
EDU» policarac(D)
ans =
la matriz no es de tamaño 2x2.
EDU»
```

En el apartado 7.2.5 vimos cómo con MATLAB es posible realizar manualmente las operaciones de renglón $R_i \leftrightarrow R_j$, $R_i \leftrightarrow \alpha R_j$ y $R_i \leftrightarrow aR_i + bR_j$; lo que resulta muy laborioso hacer cada vez que se necesite. Los siguientes listados contienen la programación de tres funciones para este fin; con ellas será más sencillo ejecutar operaciones de renglón; sin embargo, más adelante vamos a construir un programa en MATLAB que usará estas funciones y que hará este proceso mucho más amigable.

Función **intercambiofilas**

```
function m=intercambiofilas(A,i,j)
%intercambiofilas(A,i,j)
%intercambia la fila Ri con la fila Rj
%de la matriz A.
I=A(i,:);
J=A(j,:);
A(i,:)=J;
A(j,:)=I;
m=A;
```

Función **cambio_de_escala**

```
function m=cambio_de_escala(A,alpha,i)
%cambio_de_escala(A,alpha,i)
%multiplica la fila i de A
%por el escalar alpha
A(i,:)=alpha*A(i,:);
m=A;
```

Función **sumafilas**

```
%sumafilas(matriz,filai,filaj,alpha,beta)
%realiza la operacion de renglón
%R_i<--->alpha*R_i+beta*R_j}
a la matriz A
es decir, cambia la filai por alpha veces filai más
```

```
beta veces filaj
matriz(filai,:)=alpha*matriz(filai,:)+beta*matriz(filaj,:);
M=matriz;
```

Recuerde que una vez que se hayan escrito cada una de estas funciones en el editor de texto de MATLAB, se deben guardar con su respectivo nombre y con extensión `m`. Veamos ahora cómo trabajan en MATLAB:

```
EDU» M=[3 1 -1 2; 1 3 -1 1; -1 1 1 1]
M =
      3            1            -1            2
      1            3            -1            1
     -1            1             1            1
EDU»  M1=intercambiofilas(M1,1,2)
M1 =
      1            3            -1            1
      3            1            -1            2
     -1            1             1            1
EDU» M1=sumafilas(M1,2,1,1,-3)
M1 =
      1            3            -1            1
      0           -8             2           -1
     -1            1             1            1
EDU»  M1=sumafilas(M1,3,1,1,1)
M1 =
      1            3            -1            1
      0           -8             2           -1
      0            4             0            2

EDU»  M1=sumafilas(M1,3,2,2,1)
M1 =
      1            3            -1            1
      0           -8             2           -1
      0            0             2            3

EDU»  M1=cambio_de_escala(M1,1/2,3)
M1 =
      1            3            -1            1
      0           -8             2           -1
      0            0             1            3/2
```

Note que aún se están haciendo los cálculos en el formato `rat` de MATLAB. Aunque lo precedente ya es de bastante ayuda para facilitar los cálculos, todavía resulta muy tedioso tener que teclear tanta información. En el siguiente apartado diseñaremos un programa que hará mucho más amigable este proceso.

Programa Gauss_Jordan con operaciones de renglón seleccionadas por el usuario

Esta sección incluye un programa muy elemental para llevar una matriz a una forma escalonada o a la forma escalonada reducida con MATLAB, paso a paso y con operaciones de renglón elegidas por el usuario. Este programa es interactivo e incluye una opción para cancelar errores del usuario, que surgen naturalmente en este tipo de cálculos, evitando comenzar de nuevo todo el proceso. Para usuarlo, el

lector tiene que reproducir el texto del programa, salvar el archivo con extensión m, ejecutarlo en MATLAB tecleando el nombre con el que lo guardó y oprimiendo la tecla ENTER. Sobra decir que el lector es libre de usarlo y modificarlo como le convenga. La figura 7-43 contiene el guión del programa Gauss_Jordan; en él se utilizarán las funciones intercambiofilas, cambio_de_escala y sumafilas que vimos en el apartado precedente; así que el lector interesado en usar este programa tiene también que teclear esas funciones y guardarlas en el mismo directorio que el programa Gauss_Jordan. A continuación se describe *grosso modo* este programa:

```
%Programa Gauss_Jordan.
%Este programa ejecuta las operaciones de renglón
%que el usuario desee aplicar para llevar una matriz
%a una forma escalonada o a la forma escalonada reducida.
%Las operaciones intercambio de filas: R_i<--->R_j,
%cambio de escala: R_i<--->kR_i
%y suma de filas: R_i<--->aR_i+bR_j,
%se llevan a cabo pidiendo al usuario que introduzca
%la información necesaria en forma interactiva.
%Para iniciar este programa basta teclear su nombre
%en la ventana de matlab, oprimir la tecla enter
%y seguir las indicaciones.
clc;clear;
fprintf('\n         Autor: Juan Carlos Del Valle Sotelo');
fprintf('\n  ***Método de Gauss-Jordan con operaciones de renglón');
fprintf('\n           definidas por el usuario***\n\n');

format rat;
M=input('Escriba la matriz con la que desea trabajar:   ')
H=M;
f=1;
while f==1
respuesta1=input...
('\nElija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  ');
   if respuesta1==1
      respaldo=H;
      x=input('\nR_i<--->R_j: escriba la matriz [i j] :  ');
      H=intercambiofilas(H,x(1),x(2))
   elseif respuesta1==2
          respaldo=H;
          v=input('\n R_i<--->kR_i: escriba la matriz [k i]: ');
          H=cambio_de_escala(H,v(1),v(2))
   elseif respuesta1==3
          respaldo=H;
          w=input('\nR_i<-->aR_i+bR_j: escriba la matriz [i j a b]:  ');
          H=sumafilas(H,w(1),w(2),w(3),w(4))
   end
   d=input('CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  ');

   if d==0
      H=respaldo
      elseif d==1
             f=1;
      elseif d==2
             f=0;
   end
end
```

Figura 7-43 • Programa en MATLAB para aplicar el método de Gauss-Jordan paso a paso con operaciones de renglón definidas por el usuario.

1. El programa establece en el inicio el formato `rat` para trabajar con números racionales.
2. La instrucción `input` escribe un mensaje de texto en la pantalla de MATLAB y asigna a la variable que está a la izquierda del signo = el valor que el usuario teclee después de los dos puntos.
3. El ciclo `while` repite todas las instrucciones que están por debajo de esta palabra en tanto que la afirmación lógica `f==1` sea verdadera.
4. El comando de decisión `if` ya se explicó antes.
5. El comando `elseif` es similar al comando de decisión `else`, que se describió antes, pero ejecuta las instrucciones que están por debajo de él si la condición que está a su derecha es verdadera.
6. En el programa se utilizan las funciones ya programadas `intercambiofilas` , `sumafilas` y `cambio_de_escala`.
7. La opción `CANCELAR` tiene el efecto de deshacer una operación de renglón que el usuario haya realizado y considere no adecuada, regresando el programa a la matriz del paso inmediato anterior.
8. El programa debe guardarse con algún nombre y con extensión `m` en el mismo directorio donde se encuentran las funciones mencionadas antes. El nombre con el que se guardó en este caso fue `Gauss_Jordan.m` .

❍ **Nota 7.7** Los tres puntos suspensivos después de la instrucción `input` que están por debajo del comando `while`, se escribieron así únicamente para que el texto no se saliera de la caja que es visible para este libro; sin embargo, los tres puntos suspensivos se utilizan para indicar, en la sintaxis de programación en MATLAB, que la instrucción continúa en la siguiente línea. Escribir así evita que en MATLAB se genere un error si la instrucción se corta para continuar en la siguiente línea. Sin embargo, esto debe hacerse con cuidado porque una instrucción no se puede interrumpir de esta manera en cualquier parte de la misma.

Ilustremos cómo corre este programa:

```
EDU» Gauss_Jordan
                   Autor: Juan Carlos Del Valle Sotelo
            ***Método de Gauss-Jordan con operaciones de renglón
                          definidas por el usuario***

Escriba la matriz con la que desea trabajar: [3 -1 2 1; 1 2 -1 1; 2 -1 1 3]

M =
      3          -1            2            1
      1           2           -1            1
      2          -1            1            3
Elija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  1
R_i<--->R_j: escriba la matriz [i j] :  [2 3]
H =
      3          -1            2            1
      2          -1            1            3
      1           2           -1            1
CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  0

H =
      3          -1            2            1
      1           2           -1            1
      2          -1            1            3

Elija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  1
R_i<--->R_j: escriba la matriz [i j] :  [1 2]
H =
      1           2           -1            1
      3          -1            2            1
      2          -1            1            3
CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  1
```

```
Elija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  3
R_i<-->aR_i+bR_j: escriba la matriz [i j a b]:  [2 1 1 -3]
H =
      1            2           -1            1
      0           -7            5           -2
      2           -1            1            3
CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  1
Elija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  3
R_i<-->aR_i+bR_j: escriba la matriz [i j a b]:  [3 1 1 -2]
H =
      1            2           -1            1
      0           -7            5           -2
      0           -5            3            1
CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  1
Elija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  3
R_i<-->aR_i+bR_j: escriba la matriz [i j a b]:  [3 2 -7 5]
H =
      1            2           -1            1
      0           -7            5           -2
      0            0            4          -17
CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  1
Elija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  2
 R_i<--->kR_i: escriba la matriz [k i]: [1/4 3]
H =
      1            2           -1            1
      0           -7            5           -2
      0            0            1         -17/4
CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  1
Elija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  3
R_i<-->aR_i+bR_j: escriba la matriz [i j a b]:  [2 3 1 -5]
H =
      1            2           -1            1
      0           -7            0          77/4
      0            0            1         -17/4
CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  1
Elija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  3
R_i<-->aR_i+bR_j: escriba la matriz [i j a b]:  [1 3 1 1]
H =
      1            2            0         -13/4
      0           -7            0          77/4
      0            0            1         -17/4
CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  1
Elija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  2
 R_i<--->kR_i: escriba la matriz [k i]: [-1/7 2]
H =
      1            2            0         -13/4
      0            1            0         -11/4
      0            0            1         -17/4
CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  1
Elija: 1 intercambio de filas, 2 cambio de escala, 3 suma de filas:  3
R_i<-->aR_i+bR_j: escriba la matriz [i j a b]:  [1 2 1 -2]
H =
      1            0            0           9/4
      0            1            0         -11/4
      0            0            1         -17/4
CANCELAR oprimir 0. CONTINUAR oprimir 1. TERMINAR oprimir 2:  2
EDU»
```

MATLAB tiene una serie de funciones especialmente diseñadas para álgebra lineal y matrices; a su descripción nos abocamos en lo que sigue.

7.2.7 Traza, determinante, rango, inversa y transpuesta

Las funciones de MATLAB `trace`, `det`, `rank`, `inv` y `transpose` calculan, respectivamente, la traza, el determinante, el rango, la inversa y la transpuesta de una matriz. Si `A` es una matriz invertible, `inv(A)` y `A^(-1)` producen el mismo efecto, calcular la inversa de la matriz `A`; mientras que con `M'` se obtiene el mismo resultado que con `transpose(M)`, la transpuesta de la matriz `M` .

```
EDU» A=[-9 2 1 3 5; 7 -9 2 3 1; 2 1 1 4 7; -1 1 1 2 3],...
     B=[-1 2 1 2;3 1 2 1;1 1 1 1; -3 0 1 2],...
     C=[1 2 1; 1 -2 3;3 2 5]
A =
     -9            2            1            3            5
      7           -9            2            3            1
      2            1            1            4            7
     -1            1            1            2            3
B =
     -1            2            1            2
      3            1            2            1
      1            1            1            1
     -3            0            1            2
C =
      1            2            1
      1           -2            3
      3            2            5
EDU» trace(A), trace(B)

ans =
    -15
ans =
      3
EDU» [rank(A),rank(B),rank(C)]
ans =
      4            4            2
EDU» det(B)
ans =
      2
EDU» det(C)
ans =
      0
EDU» inv(B)
ans =
     -1           -1            3            0
     3/2           1           -3          -1/2
      2            3           -7            0
    -5/2          -3            8           1/2
```

Si en lugar de emplear la instrucción `inv(B)`, se ejecuta `B^(-1)`, se obtiene el mismo efecto.

```
EDU» B^(-1)
ans =
     -1          -1          3          0
     3/2          1         -3        -1/2
      2           3         -7          0
    -5/2         -3          8         1/2

EDU» M=[1 2 3; 4 5 6]
M =
     1     2     3
     4     5     6

EDU» N=transpose(M)
N =
     1     4
     2     5
     3     6

EDU» N'
ans =
     1     2     3
     4     5     6
```

Como se mencionó antes, en MATLAB no es necesario hacer ninguna declaración especial de variables para trabajar con números complejos:

```
EDU»  F=[1+i 2+3i 1; -1 -2i 1+3i; 1 -1 5i]
F =
      1    +    1i          2     +    3i          1
     -1                     0     -    2i          1     +    3i
      1                    -1                      0     +    5i
EDU» det(F)
ans =
    -13      +   35i
EDU» inv(F)
ans =
   -19/697   -  212/697i   -266/697   -  180/697i     14/41    +    3/41i
   267/1394  -  139/1394i   253/1394  +  145/1394i    -9/82    +    1/82i
    57/1394  -   61/1394i   101/1394  -  157/1394i    -1/82    -    9/82i
EDU»
```

7.2.8 Forma escalonada reducida, solución de sistemas

MATLAB tiene un comando para obtener de manera inmediata la forma escalonada reducida equivalente a una matriz. La instrucción es `rref`.

```
EDU» A=[3 1 1 2 1; -2 2 1 1 4; -5 2 1 2 3; 7 1 -1 2 1]
A =
      3          1          1          2          1
     -2          2          1          1          4
     -5          2          1          2          3
      7          1         -1          2          1
EDU» rref(A)
ans =
      1          0          0          0       1/19
      0          1          0          0      46/19
      0          0          1          0       2/19
      0          0          0          1     -16/19
```

Entonces, para resolver un sistema de ecuaciones $A\vec{x} = \vec{b}$ con MATLAB, se escribe la matriz ampliada `[A  b]`, se lleva a forma escalonada reducida con el comando `rref` y se hace sustitución regresiva en forma manual. Resolvamos con MATLAB el sistema

$$\begin{aligned} x_1 - 2x_2 + 3x_3 - 4x_4 &= -1 \\ -x_1 + 3x_2 - 5x_3 + x_4 &= 2 \\ 2x_1 - x_2 + 2x_3 - 3x_4 &= 3 \\ 4x_1 - 5x_2 + 8x_3 - 11x_4 &= 1 \end{aligned}$$

```
EDU» A=[1 -2 3 -4; -1 3 -5 1; 2 -1 2 -3; 4 -5 8 -11]
A =
      1         -2          3         -4
     -1          3         -5          1
      2         -1          2         -3
      4         -5          8        -11

b =
     -1
      2
      3
      1
EDU» AuG=[A b]
AuG =
      1         -2          3         -4         -1
     -1          3         -5          1          2
      2         -1          2         -3          3
      4         -5          8        -11          1
EDU» rref(AuG)
ans =
      1          0          0         -3          2
      0          1          0         11          3
      0          0          1          7          1
      0          0          0          0          0
```

Al hacer sustitución regresiva se obtiene: $x_3 = 1 - 7x_4$; $x_2 = 1 + 2x_3 + 3x_4 = 3 - 11x_4$; $x_1 = -1 + 2x_2 - 3x_3 + 4x_4 = 2 + 3x_4$. Esto es,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2+3r \\ 3-11r \\ 1-7r \\ r \end{bmatrix}; r \in \mathbb{R},$$

7.2.9 Valores y vectores propios, polinomio característico

MATLAB también tiene un comando para hallar los valores propios de una matriz y los vectores propios correspondientes y para encontrar el polinomio característico:

- `eig(A)` produce un vector con los valores propios de la matriz `A`.
- `[V D]=eig(A)` produce un par de matrices: una matriz diagonal `D`$=\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$, donde los λ_i son los valores propios de la matriz `A`; y una matriz `V` con columnas $\vec{v}_i$, donde cada $\vec{v}_i$ es un vector propio correspondiente al valor propio λ_i.
- `poly(A)` genera un vector $[a_n, a_{n-1}, \ldots, a_1, a_0]$, donde las componentes a_k son los coeficientes del polinomio característico,[2] $p(\lambda)=\sum_{k=0}^{n} a_{n-k}\lambda^{n-k}$, de la matriz `A`.

Ilustremos el uso de estos comandos.

Con el formato `rat` :

```
EDU» A=[2 1 0;-1 0 1; 1 3 1];
EDU» eig(A)

ans =
     -1
      2
      2
EDU» [V D]=eig(A)

V =
 -1020/5201    -985/1393    -985/1393
   919/1562        0            0
 -1121/1429    -985/1393    -985/1393
 D =
     -1            0            0
      0            2            0
      0            0            2
EDU» poly(A)

ans =
      1           -3            0            4
```

Con el formato por defecto de MATLAB (`format`) :

```
EDU» format
EDU» eig(A)

ans =
   -1.0000
    2.0000
    2.0000
EDU» [V D]=eig(A)
```

[2] $p(\lambda)=\det(\lambda I_n - A)$.

```
V =
   -0.1961   -0.7071   -0.7071
    0.5883    0.0000    0.0000
   -0.7845   -0.7071   -0.7071
   D =
   -1.0000         0         0
         0    2.0000         0
         0         0    2.0000
EDU» M=[1 1;-1 2]
M =
     1     1
    -1     2
EDU» eig(M)
ans =
   1.5000+ 0.8660i
   1.5000- 0.8660i
```

Podemos utilizar el modo simbólico de MATLAB para calcular valores y vectores propios de la siguiente manera:

```
EDU»  eig(sym(M))
ans =
[ 3/2+1/2*i*3^(1/2)]
[ 3/2-1/2*i*3^(1/2)]
```

Y para visualizar un poco mejor el resultado, utilizamos el comando `pretty` que se describió antes.

```
EDU» pretty(ans)
ans=
                              [             1/2]
                              [3/2 + 1/2 i 3   ]
                              [                ]
                              [             1/2]
                              [3/2 - 1/2 i 3   ]
EDU» [V D]=eig(sym(M))
V =
[ 1/2-1/2*i*3^(1/2), 1/2+1/2*i*3^(1/2)]
[                1,                  1]
D =
[ 3/2+1/2*i*3^(1/2),                  0]
[                0, 3/2-1/2*i*3^(1/2)]
EDU» pretty(V)
                    [             1/2                   1/2]
                    [1/2 - 1/2 i 3       1/2 + 1/2 i 3     ]
                    [                                      ]
                    [       1                   1          ]
```

No todos los comandos de MATLAB pueden funcionar en el modo simbólico.

7.2.10 Factorización *QR* y factorización *LU*

Para hallar (numéricamente) la factorización *QR* de una matriz `S`, se utiliza el comando `qr(S,0)`:

```
EDU» S=[1 1 1;0 1 -1;1 1 0;0 0 1]

S =
     1     1     1
     0     1    -1
     1     1     0
     0     0     1
EDU» [Q R]=qr(S,0)

Q =
   -0.7071    0.0000    0.4082
         0   -1.0000    0.0000
   -0.7071    0.0000   -0.4082
         0         0    0.8165
R =
   -1.4142   -1.4142   -0.7071
         0   -1.0000    1.0000
         0         0    1.2247
```

(Cfr. ejemplo 4.33)

La instrucción `[L U P]=lu(X)` produce una matriz triangular inferior `L`, una matriz triangular superior `U` y una matriz de permutación `P` tal que `P*X=L*U`:

```
EDU» X=[1 3 -1; 2 8 4; -1 3 4]

X =
     1     3    -1
     2     8     4
    -1     3     4
EDU» [L U P]=lu(X)

L =
    1.0000         0         0
   -0.5000    1.0000         0
    0.5000   -0.1429    1.0000
U =
    2.0000    8.0000    4.0000
         0    7.0000    6.0000
         0         0   -2.1429
P =
     0     1     0
     0     0     1
     1     0     0
EDU» P*X
ans =
     2     8     4
    -1     3     4
     1     3    -1
EDU» L*U

ans =
     2     8     4
    -1     3     4
     1     3    -1

EDU»
```

7.3 Excel, la herramienta Solver y programación lineal

En esta breve sección describimos el uso de una herramienta tecnológica para resolver problemas de programación lineal. No podremos tratar en toda su extensión la capacidad que tiene porque se escapa de los objetivos y alcance de este libro. El propósito es que el lector tenga un auxiliar electrónico para resolver numéricamente los problemas de programación lineal que, cuando tienen más de dos variables, pueden ser muy laboriosos y conducir eventualmente a errores cuando se hace el proceso en forma manual o que, simplemente, son virtualmente imposibles de resolver sin la ayuda de una computadora. Este auxiliar electrónico es la herramienta Solver de Excel, la hoja de cálculo con la que cuentan todas las computadoras que operan con Windows y que, por tanto, si el lector tiene a su alcance uno de estos equipos, podrá hacer uso de esta potente herramienta sin tener que desembolsar un centavo.

7.3.1 *Activación de Solver en Excel*

Vamos a trabajar con la herramienta Solver de Microsoft Office Excel 2003. La razón de esta elección se debe únicamente al hecho de que fue más sencillo capturar las imágenes que aparecen más adelante en esa versión, que en la versión 2007 de esta hoja de cálculo. Explicaremos cómo se carga y activa esta herramienta en cada una de las dos versiones y después daremos la descripción de su uso en la versión 2003 solamente; ya que no existe ningún cambio que merezca mencionarse en la versión 2007 para los objetivos de esta sección.

Versión 2003

Activación Para activar la herramienta Solver en Microsoft Office Excel 2003, hay que hacer clic en el menú Herramientas de Excel y después en la opción Solver de ese menú como se ilustra en la figura 7-44, si es que esta opción aparece en el menú Herramientas. Si la opción Solver no aparece en el menú Herramientas de Excel, es que no está cargada; en el siguiente ítem se indica lo que se tiene que hacer en ese caso.

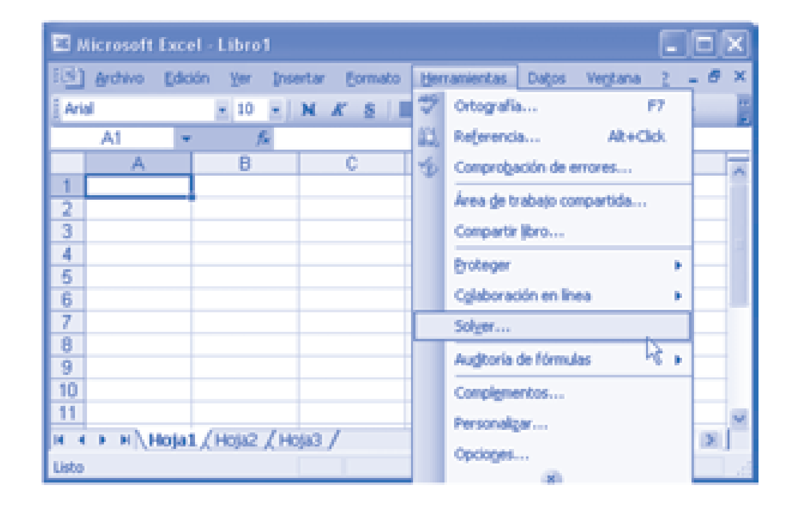


Figura 7-44 •

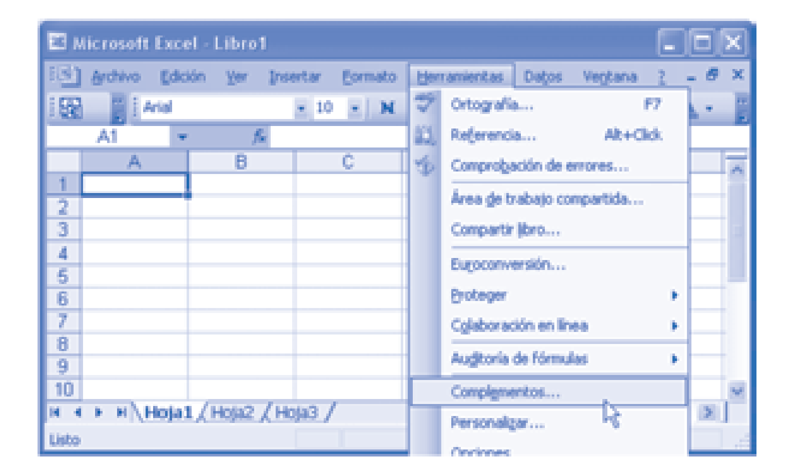


Figura 7-45 •

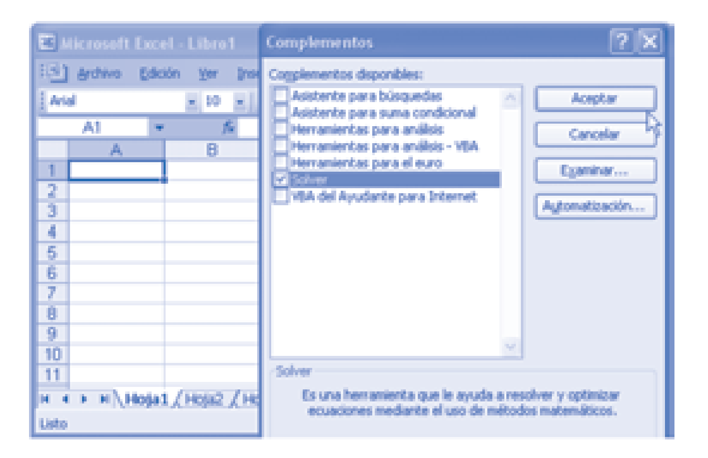


Figura 7-46 •

Carga Para cargar la herramienta Solver hay que hacer clic en el menú Herramientas y después en la opción Complementos, como se ilustra en la figura 7-45, entonces se desplegará la ventana mostrada en la figura 7-46. Luego se coloca el cursor en el pequeño cuadrado a la izquierda de la opción Solver, se hace clic en él y después en el botón Aceptar del mismo menú en esa ventana, figura 7-46. De esta manera, la herramienta Solver quedará cargada en el menú Herramientas y Solver podrá activarse como se indica en el precedente ítem.[3]

Versión 2007

Activación Para activar la herramienta Solver en Microsoft Office Excel 2007, hay que hacer clic en el menú Datos de Excel y después en la opción Solver de ese menú, si es que esta opción aparece en el menú Datos. Si la opción Solver no aparece en el menú Datos de Excel, es que no está cargada; en el siguiente ítem se indica lo que se tiene que hacer en ese caso.

Carga Para cargar la herramienta Solver en Microsoft Office Excel 2007:

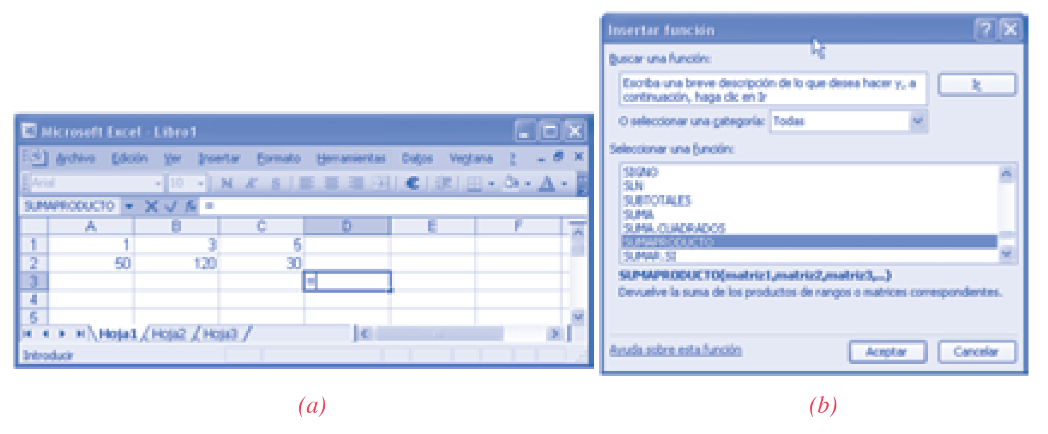
[3]Es posible que este procedimiento no pueda llevarse a cabo; en tal caso, después de haber hecho clic en la opción Aceptar, se desplegará un mensaje indicando que se requiere insertar el disco de instalación de Microsoft Office 2003 en la computadora; inserte el disco y siga las instrucciones.

1. Hacer clic en el botón de Microsoft Office y después hacer clic en `Opciones de Excel`.
2. Hacer clic en `Complementos` y, en el cuadro `Administrar`, seleccionar `Complementos de Excel`.
3. Hacer clic en `Ir`.
4. En el cuadro `Complementos disponibles`, activar la casilla de verificación `Complemento Solver` y hacer clic en `Aceptar`.

De esta manera, la herramienta Solver quedará cargada en el menú Datos y Solver podrá activarse como se indica en el ítem activación.

7.3.2 La función SUMAPRODUCTO de Excel

Para usar Solver necesitamos de la función `SUMAPRODUCTO`, la cual multiplica elemento a elemento las componentes de dos filas de una hoja de cálculo que tengan el mismo número de componentes y suma los resultados. Supongamos que deseamos aplicar esta función a la primera fila, de las celdas `A1` a `C1`, y a la segunda fila, de las celdas `A2` a `C2`, y que el resultado quede registrado en la celda `D3` de la hoja de cálculo mostrada en la figura 7-47*(a)*. Entonces, se hace clic en esta celda y después en el botón f_x de inserción de fórmulas;[4] inmediatamente Excel desplegará la ventana Insertar función que se muestra en la figura 7-47*(b)*. En ella se selecciona la categoría `Todas` (utilizando la barra de desplazamiento vertical y haciendo clic sobre esa opción) del menú o seleccionar una categoría y con la barra de des-

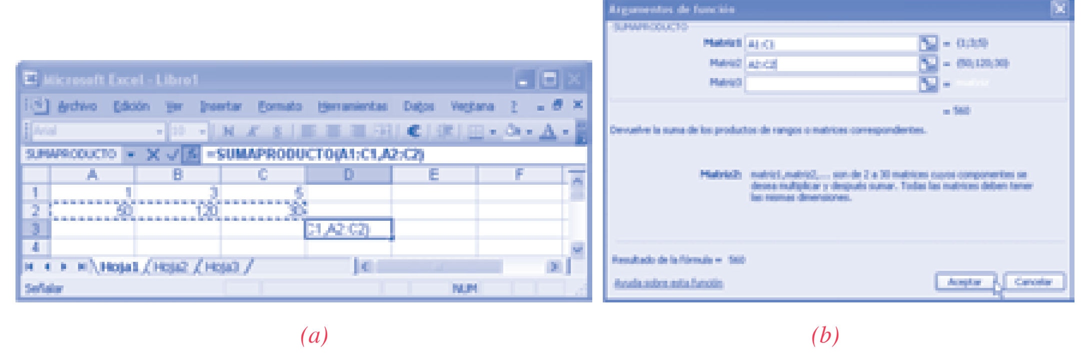


(a) *(b)*

Figura 7-47 •

plazamiento se busca la función `SUMAPRODUCTO` del menú Seleccionar una función. Se hace clic sobre esta opción y en Aceptar de esa ventana, como se ilustra en la misma figura (7-47)*(b)*. Después de esto Excel mostrará una nueva ventana, figura 7-48*(b)*. Hay que hacer clic en el espacio que está después de la palabra `Matriz1`, oprimir el botón izquierdo del ratón de la computadora en la celda `A1` y, sin dejar de oprimir ese botón, arrastrar el ratón hasta la celda `C1` y dejar de oprimir; se repiten estas dos operaciones ahora con el espacio que está después de la palabra `Matriz2` y con las celdas `A2` a `C2` de la segunda fila de la hoja de cálculo (cfr. figura 7-48*(a)* y *(b)*). De esta manera quedarán registradas esas celdas en el menú de la función `SUMAPRODUCTO` como se ilustra en la figura 7-48*(b)*. Por último se

[4] En la versión 2007 este botón aparece en el menú Fórmulas.



(a) (b)

Figura 7-48 •

hace clic en Aceptar de ese menú, obteniéndose la hoja de cálculo mostrada en la figura 7-49. Se ha creado así una fórmula dinámica en la celda D3 y cada vez que se cambien los valores en las celdas A1-C1 y A2-C2, se obtendrá en forma automática el valor correspondiente de la función SUMAPRODUCTO, aplicada a esos valores en la celda D3.

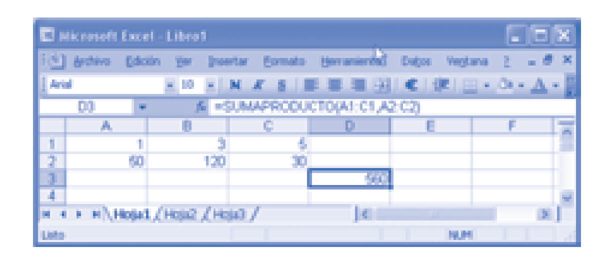


Figura 7-49 •

7.3.3 Resolución de problemas de programación lineal con Solver

En este apartado utilizaremos el problema de programación lineal que dimos en el ejemplo 6.21 (pág. 613) para ilustrar el uso de Solver en la solución de problemas de programación lineal. Repetimos aquí la información por comodidad en la lectura:

Una compañía tiene dos talleres donde produce cuatro tipos de libreros para armar que se venden en tiendas de autoservicio. La siguiente tabla contiene la información relativa a los datos de producción, costos y demanda. Encontrar el número de días que tiene que trabajar cada taller durante seis meses para proporcionar, de la manera más económica, los libreros requeridos.

Tipo de librero	Producción diaria taller 1	Producción diaria taller 2	Demanda semestral
A	120	40	4 800
B	80	80	6 400
C	40	120	4 800
D	60	100	6 000
Costos diarios	$40 000	$60 000	

(7.1)

Recordemos que se tiene que resolver el problema de programación lineal:

Hallar el mínimo de

$$C = 40\,000x + 60\,000y$$

sujeto a las restricciones

$$\begin{aligned} 120x + 40y &\geq 4\,800 \\ 80x + 80y &\geq 6\,400 \\ 40x + 120y &\geq 4\,800 \\ 60x + 100y &\geq 6\,000 \\ x, y &\geq 0 \end{aligned} \tag{7.2}$$

Preparación de la hoja de cálculo que utilizará Solver

Con la información de la tabla anterior vamos a construir la hoja de cálculo que se muestra en la figura 7-50. En ella hemos colocado algunos textos para facilitar la comprensión de la misma; obviamente los textos son opcionales y no necesarios para usar Solver. Toda la información numérica de la hoja coincide en forma respectiva con la de la tabla 7.1; excepto las cantidades que están por debajo de los textos `Días taller 1` y `Días taller 2`. Estas dos cantidades representan información numérica inicial que hemos colocado para utilizar Solver en la solución de este problema; ambos números en realidad son arbitrarios y pudimos colocar cualquier otro par. También hemos añadido una columna con un encabezado más, `Cant. Libreros`, y una celda que contiene el enunciado `Costo total` y los signos `>=` en las celdas `E4` a `E7`; los cuales se consideran también texto auxiliar. La demás información de la hoja de cálculo se explica por sí sola observando la tabla 7.1.

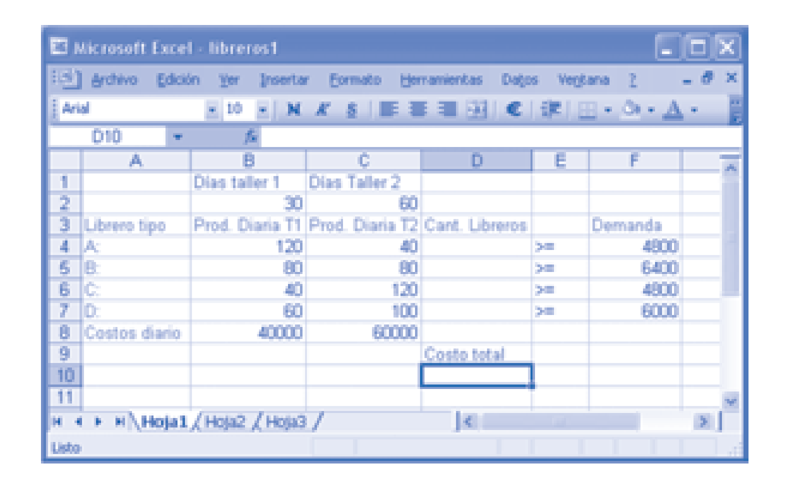


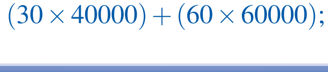
Figura 7-50 •

Las celdas `B2` y `C2` contienen el número de días que proponemos trabajen los talleres 1 y 2, respectivamente, como lo indican los textos por encima de esas celdas; en el caso de la figura 7-50, 30 y 60. Con esas dos celdas, `B2` y `C2`, y las celdas `B4` y `C4`, creamos una fórmula utilizando la función `SUMAPRODUCTO`, que describimos en el apartado anterior, para que en la celda `D4` se obtenga la cantidad total de libreros de tipo `A` que producen los dos talleres

$$(30 \times 120) + (60 \times 40);$$

Figura 7-51 •

pero esta vez con una pequeña modificación: después de selecionar las celdas B2 y C2 como argumentos en Matriz1, se oprime la tecla F4 con el objetivo de fijar esas celdas en Matriz1; la figura 7-51 ilustra ese proceso. Concluido este paso, en la celda D4 deberá aparecer el valor 6000. Ahora hay que crear, con la función SUMAPRODUCTO, fórmulas en las celdas D5, D6 y D7 para obtener la cantidad total de libreros tipos B, C y D, respectivamente; para ello es suficiente copiar la fórmula de la celda D4: se hace clic en la celda D4 con el botón izquierdo del ratón y sin dejar de oprimir este botón se arrastra el ratón hasta la celda D7, como se indica en la figura 7-52 y se suelta ese botón. Después hay que crear una fórmula, utilizando la función SUMAPRODUCTO, en la celda D10 para obtener el costo total



$$(30 \times 40000) + (60 \times 60000);$$

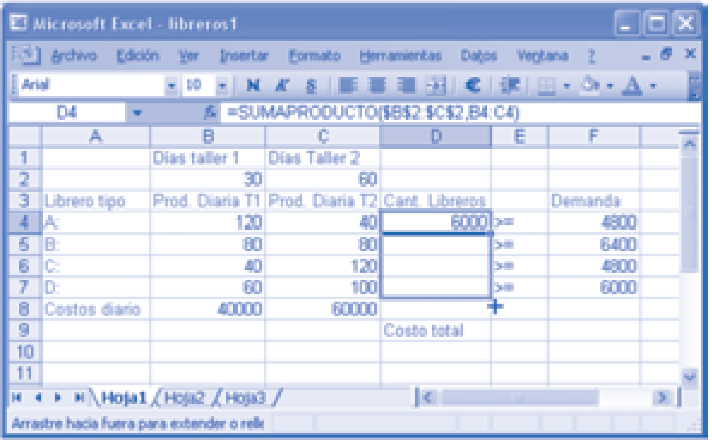
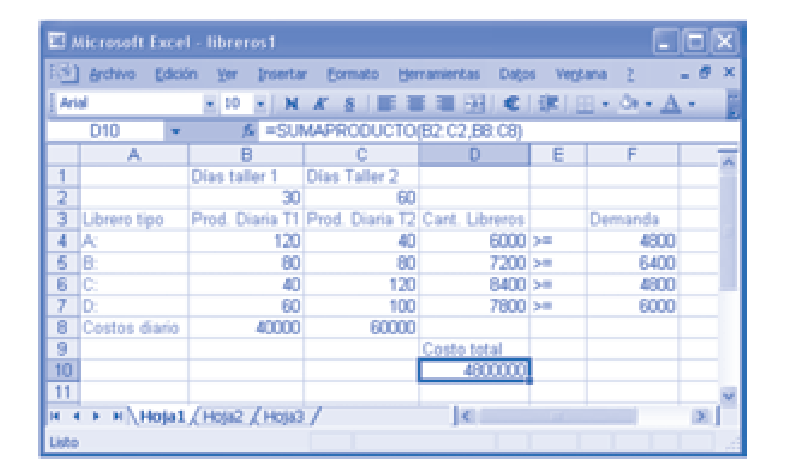


Figura 7-52 •

utilizando las celdas B2, C2 y B8, C8 en los argumentos Matriz1 y Matriz2, respectivamente, de la función SUMAPRODUCTO. Al concluir este proceso se obtendrá la hoja de cálculo contenida en la figura 7-53. Los valores por debajo de Cant. Libreros son las cantidades totales de libreros tipos A, B, C y D que fabrican ambos talleres; mientras que la celda D10 contiene el costo total de producción; todo ello si el primer taller trabaja 30 días y el segundo 60 días. Ya que hemos creado fórmulas para cada una de estas celdas, si se cambian los valores 30 y 60 de las celdas B2 y C2, se obtendrán inmediatamente los valores correspondientes de cantidades de libreros y costo total. El lector puede verificar con su hoja de cálculo este hecho modificando valores. Es claro que también se pueden cambiar los demás parámetros si así se requiere; por ejemplo, la producción de determinado tipo de librero



	A	B	C	D	E	F
1		Días taller 1	Días Taller 2			
2		30	60			
3	Librero tipo	Prod. Diaria T1	Prod. Diaria T2	Cant. Libreros		Demanda
4	A:	120	40	6000	>=	4800
5	B:	80	80	7200	>=	6400
6	C:	40	120	8400	>=	4800
7	D:	60	100	7800	>=	6000
8	Costos diario	40000	60000			
9				Costo total		
10				4800000		
11						

Figura 7-53 •

de algún taller o el costo diario de alguno de los talleres, etc. De esta manera, la hoja de cálculo ya está preparada para utilizar Solver y resolver el problema.

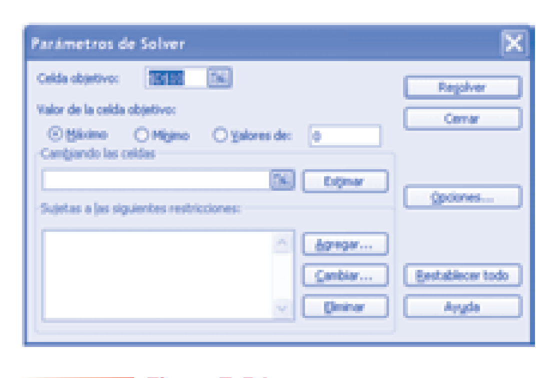


Figura 7-54 •

Utilización de Solver

En la hoja de cálculo que preparamos en el segmento anterior, se hace clic sobre la celda que contiene el costo total, celda `D10`, y se activa Solver como se describió en el apartado 7.3.1. Entonces, se abrirá la ventana Parámetros de Solver contenida en la figura 7-54. En ella, puesto que colocamos antes el cursor en la celda `D10` e hicimos clic en ésta, tiene seleccionada esa celda en la opción `Celda objetivo`; si no es así, hay que hacer clic en `D10` de la hoja de cálculo para seleccionar ésta como celda objetivo en Solver. Ahora, en la ventana Parámetros de Solver hay que seleccionar la opción `Mínimo`, haciendo clic en el botón que está a la izquierda de este texto (cfr. figura 7-55*(b)*) —porque en este ejemplo se tiene, en particular, un problema de minimización; en caso contrario se debe elegir la opción `Máximo`. Luego se hace clic en la casilla que está por debajo del texto `Cambiando las celdas` y se seleccionan las celdas `B2` a `C2` de la hoja de cálculo; para ello hay que colocar el cursor en la casilla `B2`, oprimir el botón izquierdo del ratón y, sin dejar de oprimir ese

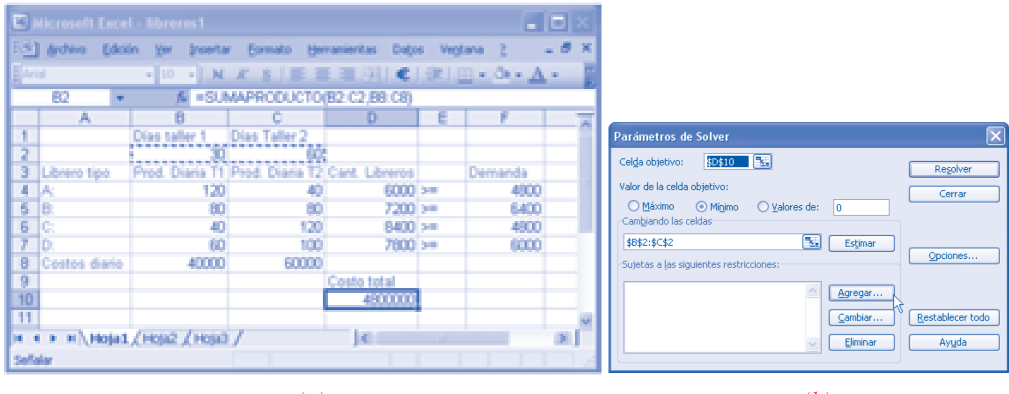


Figura 7-55 • *(a)* *(b)*

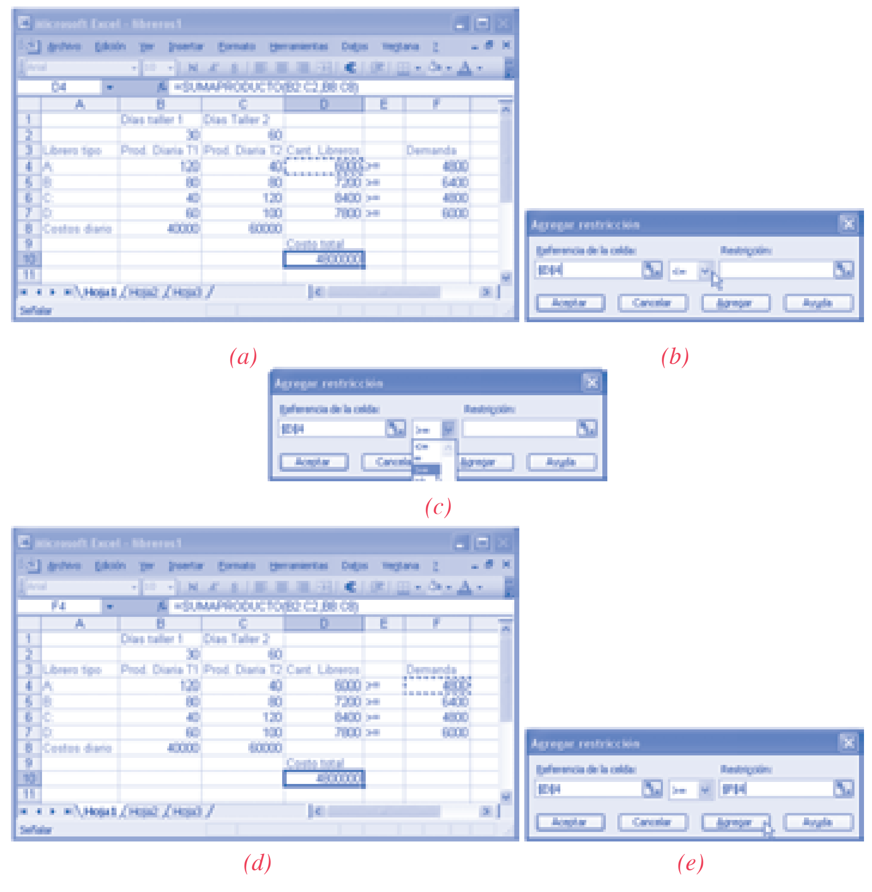


Figura 7-56 •

botón, arrastrar el ratón hasta la celda C2 y soltar el botón izquierdo, figura 7-55*(a)*. Después de esto hay que hacer clic en Agregar del menú Parámetros de Solver para capturar las restricciones en la casilla que está por debajo del texto Sujeta a las siguientes restricciones:, como se ilustra en la figura 7-55*(b)*. Entonces, se desplegará la ventana Agregar restricción contenida en la figura 7-56*(b)*. Se hace clic en la casilla Referencia de la celda y después en la celda D4 de la hoja de cálculo como se hace patente en la figura 7-56*(a)*. Luego, se elige la opción >= de ese menú, figura 7-56*(c)*; se hace clic en la celda F4 de la hoja de cálculo para elegir esta celda en la opción Restricción, como se ilustra en las figuras 7-56*(d)* y *(e)*. Se elige la opción Agregar de ese menú y se repite el proceso con las celdas D5-F5, D6-F6 y D7-F7. Cuando se haya llegado a este último paso, se elige la opción Aceptar del menú Agregar restricción y entonces se desplegará la ventana mostrada en la figura 7-57*(b)*. Este menú contiene, en la casilla Sujeta a las restricciones, el equivalente para Solver de las restricciones 7.2 del problema de los libreros y en la casilla Celda objetivo la función de costo de este problema; pero con valores numéricos concretos, los contenidos en la casilla Cambiando las celdas. Es prudente en este momento comprobar que toda la información capturada en el menú Parámetros de Solver (figura 7-57*(b)*) es correcta, comparando con la hoja de cálculo (figura 7-57*(a)*). De ser así, para continuar con el proceso, hay que seleccionar Opciones en este menú haciendo clic en Opciones;[5] entonces se desplegará la ventana

[5]En caso contrario se debe colocar el cursor en la restricción que se desee modificar, hacer clic en ella, seleccionar la opción Cambiar y hacer las modificaciones pertinentes o, según el caso, utilizar las opciones Agregar o Eliminar.

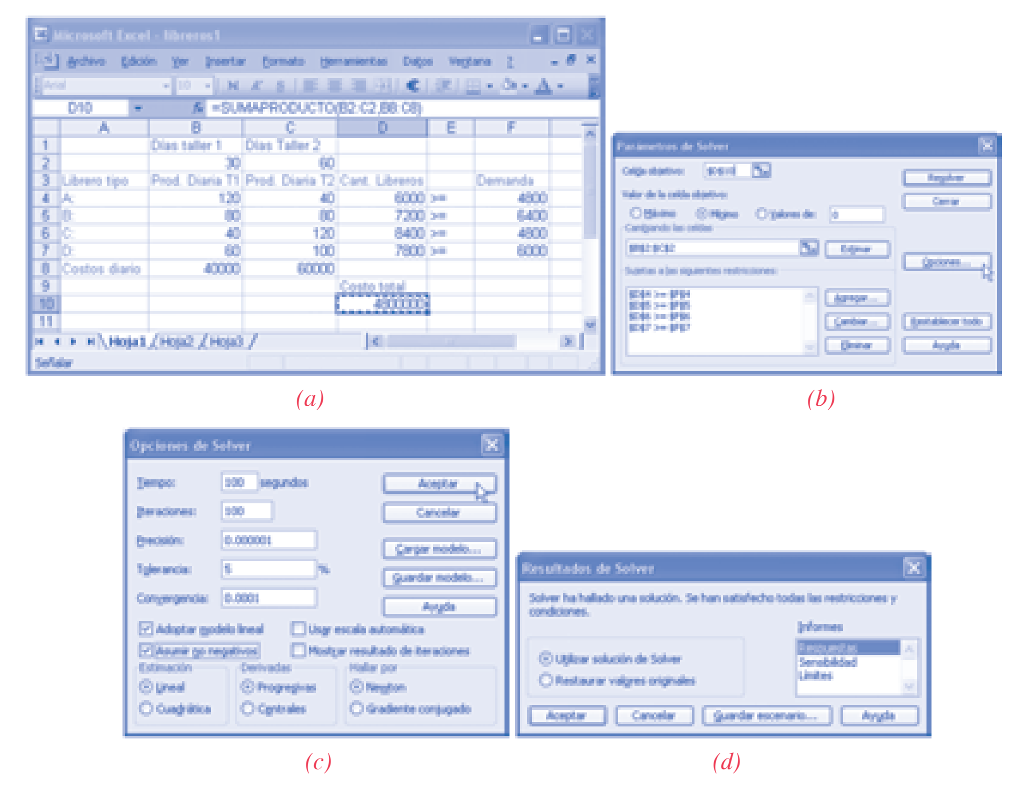


Figura 7-57 •

que contiene la figura 7-57*(c)*. En este menú se deben seleccionar las opciones `Adoptar modelo lineal` y `Asumir no negativos` haciendo clic en los recuadros que están a la izquierda de ambas opciones; después se debe hacer clic en `Aceptar` de ese menú como se indica en la figura 7-57*(c)*. Así, Excel volverá a mostrar la ventana con el menú que contiene la figura 7-57*(b)* y en ese menú hay que oprimir `Resolver`; inmediatamente Excel desplegará una nueva ventana, contenida en la figura 7-57*(d)*, con el menú Resultados de Solver. En este menú se selecciona la opción `Utilizar solución de Solver`, haciendo clic en el botón a la izquierda, y en el submenú `Informes`, se pueden seleccionar una o todas las opciones: `Respuestas`, `Sensibilidad` o `Límites` haciendo clic en cualquiera de ellas. En nuestro caso hemos seleccionado la primera; pues es la única que describiremos en este libro. Por último se hace clic en `Aceptar` de este menú; entonces en nuestra hoja de cálculo se mostrarán los resultados que la herramienta Solver ha calculado en las celdas `B2`, `C2` y `D10` como se ilustra en la figura 7-58. Esto es, el taller 1 debe trabajar `50` días y el taller 2 debe trabajar `30` días con un costo mínimo de `$3 800 000`; que es la misma respuesta que encontramos resolviendo geométricamente este problema.

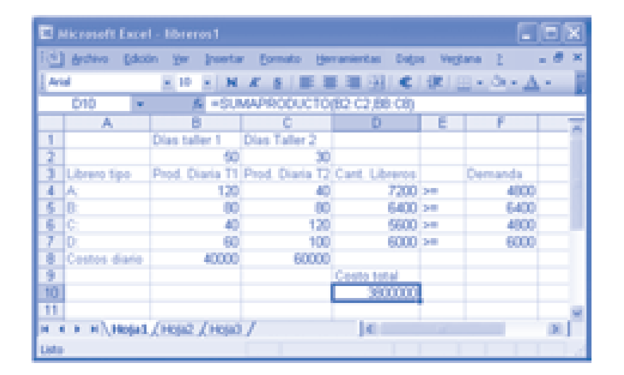


Figura 7-58 •

El informe `Respuestas`

Puesto que en el menú `Informes` elegimos la opción `Respuestas`, la herramienta Solver generará el informe contenido en la hoja `Informe de respuestas 1` del libro que contiene nuestra hoja de cálculo en Excel (cfr. figura 7-58); al hacer clic en la pestaña con este nombre se desplegará la hoja de cálculo contenida en la figura 7-59. El informe `Respuestas` consta de tres partes:

- `Celda objetivo (Mínimo)`. Que indica la celda objetivo donde se encuentra programada la fórmula para el costo del problema, su nombre, el valor inicial (original para Solver) y el valor final (para Solver), la respueta que calcula la herramienta.
- `Celdas cambiantes`. Que informa las celdas donde se encuentran los valores numéricos que van a cambiar durante el transcurso de la solución del problema, los días que trabaja cada taller, sus nombres, valores iniciales y valores finales (respuesta que da la herramienta Solver).
- `Restricciones`. Que indica las celdas donde se encuentran las restricciones, sus respectivos nombres, el tipo de restricción en cada caso, el `Estado` (opcional u obligatorio) para cada restricción y el valor de `Divergencia`. De acuerdo con Solver la primera y tercera restricciones de (7.2) (cfr. pág. 807) son opcionales y la segunda y cuarta son obligatorias. Esto significa que si se excluyen las restricciones primera y tercera y se conservan las restricciones segunda y cuarta de (7.2), entonces, la función de costo para este problema alcanza el mismo valor mínimo (3 800 000) en igual punto ((50, 30)). El lector puede verificar esta aseveración resolviendo geométricamente el mismo problema, pero sólo con las restricciones segunda y cuarta. Sin embargo, aunque matemáticamente es posible excluir las restricciones opcionales que la herramienta Solver marque como tales en el informe `Respuestas` de un problema de programación lineal, hay que tener mucho cuidado con el manejo de esta información; ya que en la práctica algunas restricciones son imposibles de excluir. Por ejemplo, el número de horas que por condiciones legales puede trabajar a lo más un obrero, el número máximo de horas que por condiciones mecánicas puede funcionar un dispositivo, la cantidad de materia prima con la que se cuenta, el espacio para almacenar, cantidad de dinero para insumos, etc., son restricciones que no se pueden excluir en un problema real. Por último, los valores de divergencia en cada restricción, son las cantidades de productos que, para este problema, sobrepasan la demanda requerida; así, se produce un excedente de 2 400 libreros del tipo *A* y un excedente de 800 libreros de tipo *D* con igual costo mínimo.

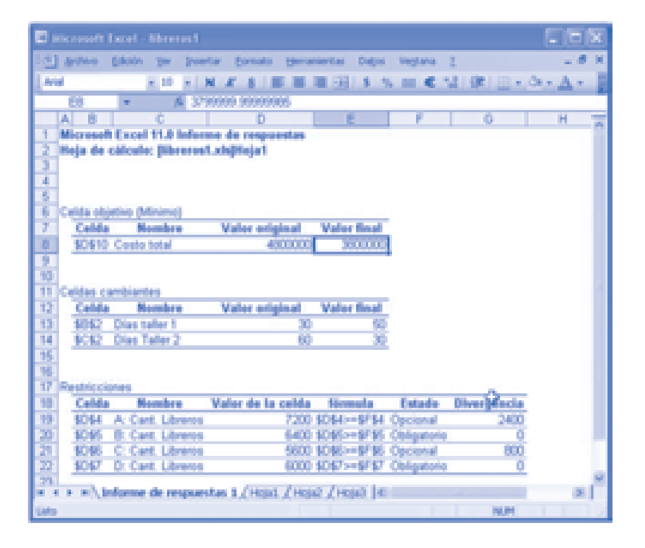


Figura 7-59 •

7.4 Ejercicios propuestos

El lector encontrará la respuesta a los ejercicios en cursiva en el apéndice E al final del libro.

La calculadora HP 50g y álgebra lineal (respuestas en páginas 1097-1098)

1 Sean $\vec{u} = (-1, \sqrt{2}, \sqrt{3}, 2)$, $\vec{v} = (2, -2\sqrt{2}, 3, -\sqrt{3})$, $\vec{w} = (-1, \sqrt{5}, 3, 0)$. Calcular en forma exacta (utilizar, para simplificar, la función SIMPLIFY del submenú REWRITE del menú CONVERT):

(a) $\sqrt{2}\vec{u} - 3\vec{v} + \sqrt{3}\vec{w}$ *(b)* $2\vec{u} \cdot \vec{v}$. *(c)* $\|\vec{u}\|$. *(d)* $\|\vec{u} - 3\vec{w}\|$. *(e)* $\|\vec{u}\| - 3\|\vec{v}\|$.

2 Hacer los cálculos del ejercicio precedente en forma numérica.

3 Sean $\vec{u} = (-1, 1, \sqrt{3})$, $\vec{v} = (-2, \sqrt{2}, \sqrt{3})$, $\vec{w} = (-1, 1, 2)$. Calcular en forma exacta:

(a) $\vec{u} \times \vec{v}$. *(b)* $\vec{u} \times (\vec{v} - \vec{w})$. *(c)* $(\vec{u} \times \vec{v}) \cdot \vec{w}$. *(d)* $(\vec{w} \times \vec{v}) \times \vec{v}$.

4 Repetir los cálculos del ejercicio precedente en forma numérica.

5 Calcular $(1, a, a^2) \cdot (-a, 1/a, a^3)$; utilizar, para simplificar, la función SIMPLIFY del submenú REWRITE del menú CONVERT.

6 Calcular $(1, a, a^2) \times (-a, 1/a, a^3)$; utilizar, para simplificar, la función SIMPLIFY del submenú REWRITE del menú CONVERT.

7 Encontrar una base para el espacio generado por los vectores $\vec{u} = (-1, 1, 2, 3)$, $\vec{v} = (1, -3, 0, 2)$, $\vec{w} = (-5, 9, 6, 5)$ y $\vec{z} = (3, 1, 1, 0)$.

8 Encontrar una base para la intersección de los espacios generados por

$$\{(-1, 2, -1, 1, 2), (1, -1, 2, 3, -1), (3, -1, 0, 1, 0)\}$$

y

$$\{(-1, 3, -1, 2, 1), (-1, 0, 1, 2, 1), (1, -4, 2, 1, -1)\}$$

9 Resolver los ejercicios 8 a 15 del capítulo 1 utilizando la calculadora HP 50g.

10 Sea $A = \begin{bmatrix} a & a & -a \\ -a & a & a \\ a & 0 & 0 \end{bmatrix}$, donde $a \neq 0$. Calcular:

(a) El rango de A. *(b)* El determinante de A. *(c)* $\|A\|$. *(d)* La inversa de A, si es que existe.

11 Sea $F = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$.

(a) Utilizar la función REF para encontrar una forma escalonada equivalente a F para verificar que sus columnas son vectores linealmente independientes de $\mathbb{R}^4$.
(b) Hallar la factorización QR de la matriz F; utilizar, para simplificar, la función SIMPLIFY del submenú REWRITE del menú CONVERT.

12 Utilizar la función REF para probar que los vectores $\vec{v}_1 = (1, -1, 2, 1)$, $\vec{v}_2 = (2, 0, 1, 1)$ y $\vec{v}_3 = (1, 1, -1, 1)$ son linealmente independientes y encontrar una base ortonormal para el subespacio $\text{gn}(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ de $\mathbb{R}^4$.

13 Utilizar la función RREF para resolver por sustitución regresiva los sistemas de los ejercicios 66, 67 y 72 del capítulo 1.

14 Resolver, utilizando la función RREF, el ejercicio 77 del capítulo 1.

15 Activar el modo STEP/STEP y encontrar la forma escalonada reducida de la matriz

$$M=\begin{bmatrix}4 & 3 & -1 & 2\\ 3 & -1 & 1 & 1\\ 5 & 1 & 2 & 7\\ 7 & -1 & 2 & 0\\ 2 & 3 & -2 & 3\end{bmatrix}.$$

16 Activar el modo STEP/STEP y encontrar la inversa de la matriz, si es que existe

$$A=\begin{bmatrix}2 & -1 & 7 & 2\\ -3 & 2 & 3 & -3\\ 5 & 3 & 2 & 4\\ 7 & 4 & 5 & 2\end{bmatrix}.$$

17 Activar el Soft MENU y emplear las funciones RCI, RCIJ y RSWP para encontrar la forma escalonada reducida de la matriz

$$A=\begin{bmatrix}3 & -2 & 7 & 9 & 1\\ 4 & -1 & 2 & 3 & 1\\ 5 & -1 & 2 & 3 & 2\\ 7 & -9 & 2 & 1 & 4\end{bmatrix}.$$

18 Utilizar la función AuM, programada en el apartado 7.1.12 y las funciones RCI, RCIJ y RSWP para hallar la inversa de la matriz, si es que existe.

$$A=\begin{bmatrix}2 & -1 & 1 & 1\\ 3 & -2 & 1 & 0\\ 4 & -1 & 1 & 3\\ 7 & -2 & 1 & 3\end{bmatrix}.$$

19 Hallar bases para la imagen y el núcleo de las transformaciones lineales cuya representación matricial, respecto a las bases canónicas, está dada por la matriz A:

(a) $A=\begin{bmatrix}1 & -1 & 2\\ 1 & 2 & -3\\ 3 & 2 & 1\\ 2 & -1 & -1\end{bmatrix}$. *(b)* $A=\begin{bmatrix}4 & -3 & 1 & 2\\ 0 & -1 & 1 & 1\\ 2 & -1 & 2 & 5\\ 7 & -1 & 2 & 7\end{bmatrix}$.

En los ejercicios 20 a 23 hallar los valores propios, vectores propios correspondientes y polinomio característico de la matriz A.

20 $A=\begin{bmatrix}-1 & 1 & 0\\ 1 & 2 & 1\\ 0 & 3 & -1\end{bmatrix}$. **21** $A=\begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix}$. **22** $A=\begin{bmatrix}2 & 1\\ -2 & 4\end{bmatrix}$. **23** $A=\begin{bmatrix}1 & -1 & 1\\ -1 & 2 & 1\\ 1 & 1 & -1\end{bmatrix}$.

En los ejercicios 24 a 31 poner en modo Complex la calculadora para calcular las matrices o valores indicados si $A=\begin{bmatrix}1-i & 2+3i\\ 4+5i & 7-i\\ 2+5i & 3-2i\end{bmatrix}$, $B=\begin{bmatrix}i & 1-i\\ 3i & 2+i\end{bmatrix}$, $C=\begin{bmatrix}-5+4i & -3+7i\\ 9+2i & -4-7i\\ 8+3i & -2+6i\end{bmatrix}$

y $D=\begin{bmatrix}i & 2+3i & -i & 2\\ 1-i & 5+3i & -2-2i & 1\\ -1+3i & 2+5i & 3i & 1-7i\\ -4 & 2 & -i & i\end{bmatrix}$.

24 $(1-i)A+2iC$. **25** $AC-(2+i)C$. **26** $\det(B)$.
27 $\det(D)$. **28** B^{-1} (si B es invertible). **29** D^{-1} (si D es invertible).
30 Los valores propios y vectores propios correspondientes de la matriz B.
31 Los valores propios y vectores propios correspondientes de la matriz D.

MATLAB y álgebra lineal (respuestas en páginas 1098-1099)

En los ejercicios 32 a 40 calcular la matriz o el número ahí indicado si

$$A=\begin{bmatrix}-4 & 1 & 3 & -1 & 2\\ 1 & -1 & 1 & 0 & 7\\ -11 & 2 & 1 & -3 & 4\\ 2 & -1 & 13 & 21 & 17\\ 11 & -1 & 3 & 1 & 4\end{bmatrix} \text{ y } B=\begin{bmatrix}-3 & 7 & -9 & 12\\ 4 & 1 & 0 & -2\\ 1 & 1 & 12 & 11\\ 21 & -4 & 3 & 2\\ -2 & 3 & 1 & 5\end{bmatrix}.$$

32 A^5+A. **33** A^3B. **34** $A^2(A^t)^2$.
35 B^tA^2. **36** $(AB)(AB)^t$. **37** $((3A)^2-2A)(3B)$.
38 $\det(A)$. **39** A^{-1} en forma aproximada. **40** A^{-1} en forma exacta.

En los ejercicios 41 a 49 calcular la matriz o el número ahí indicado si

$$A=\begin{bmatrix}1/2 & -1/3 & 1 & 1/7\\ 2/3 & 3/2 & 1/8 & -1/10\\ 2/7 & -3/5 & 1/2 & 2\\ -1 & 2 & 1 & 3\end{bmatrix} \text{ y } B=\begin{bmatrix}-1/8 & 7/5 & 1/3\\ 2 & -1 & 0\\ 1/2 & 2/3 & -1/9\\ 1 & 0 & -3/7\end{bmatrix}.$$

41 A^5+A. **42** A^3B. **43** $A^2(A^t)^2$.
44 B^tA^2. **45** $(AB)(AB)^t$. **46** $((3A)^2-2A)(3B)$.
47 $\det(A)$. **48** A^{-1} en forma exacta. **49** A^{-1} aproximada.

En los ejercicios 50 a 58 calcular la matriz o el número ahí indicado si

$$A=\begin{bmatrix}1-i & 2+3i & -1 & 5-3i\\ i & 3-7i & 8-4i & 5+2i\\ 5-3i & 4+i & -7-2i & -1+3i\\ -2+3i & 3+7i & -4+3i & 4-3i\end{bmatrix} \text{ y } B=\begin{bmatrix}2+2i & 3-i & 7-3i\\ 2-5i & -1+3i & i\\ 7-i & 3+2i & -9-6i\\ -1+5i & 7+4i & -3+7i\end{bmatrix}.$$

50 A^5+A. **51** A^3B. **52** $A^2(A^t)^2$.
53 B^tA^2. **54** $(AB)(AB)^t$. **55** $((3A)^2-2A)(3B)$.
56 $\det(A)$. **57** A^{-1}en forma aproximada. **58** A^{-1} en forma exacta.

59 Resolver utilizando MATLAB los ejercicios 8 a 15 del capítulo 1.

60 Sea $A=\begin{bmatrix}a & a & -a\\ -a & a & a\\ a & 0 & 0\end{bmatrix}$, donde $a\neq 0$. Calcular:

(*a*) El rango de A. (*b*) El determinante de A. (*c*) La inversa de A, si es que existe.

61 Calcular $\begin{vmatrix}a & b\\ c & d\end{vmatrix}$.

62 Calcular la inversa del ejercicio precedente si se supone que $ad-bc\neq 0$.

63 Hallar $\begin{vmatrix}a & b & c\\ d & e & f\\ g & h & u\end{vmatrix}$. **64** Si $A=\begin{bmatrix}0 & 0 & e\\ 1 & 0 & f\\ 0 & 1 & g\end{bmatrix}$, con $e\neq 0$: (*a*) Calcular $\det(A)$.
(*b*) Hallar A^{-1}.

En los ejercicios 65 a 69 utilizar el programa Gauss_Jordan para encontrar una matriz en forma escalonada equivalente a la matriz dada.

65 La matriz del ejercicio 58 del capítulo 1.

66 La matriz del ejercicio 60 del capítulo 1.

67 La matriz del ejercicio 61 del capítulo 1.

68 $A = \begin{bmatrix} 4 & 3 & -2 & 1 & 2 \\ 3 & -1 & 1 & 3 & 4 \\ 5 & 7 & -1 & 2 & 1 \\ 3 & -2 & 2 & 1 & 5 \\ -2 & 1 & 1 & 0 & 2 \\ 2 & -1 & 1 & 1 & 0 \end{bmatrix}$. **69** $Z = \begin{bmatrix} i & -i & 1+i & 4i \\ -i & 2-3i & 0 & 1 \\ 1+i & 2 & -1 & 0 \\ 2+3i & -2+i & 1 & 1 \\ 4-2i & 3+i & 0 & 2i \end{bmatrix}$.

70 Resolver el sistema lineal del ejercicio 75 del capítulo 1, utilizando la instrucción `rref` y haciendo sustitución regresiva.

71 Resolver el sistema lineal del ejercicio 76 del capítulo 1, utilizando la instrucción `rref` y haciendo sustitución regresiva.

72 Resolver el sistema lineal del ejercicio 114 del capítulo 1, utilizando la instrucción `rref` y haciendo sustitución regresiva.

73 Resolver el sistema lineal del ejercicio 116 del capítulo 1, utilizando la instrucción `rref` y haciendo sustitución regresiva.

74 Encontrar el rango de la matriz del ejercicio 65.

75 Encontrar el rango de la matriz del ejercicio 69.

76 Hallar bases para la imagen y el núcleo de las transformaciones lineales cuya representación matricial, respecto a las bases canónicas, está dada por la matriz A:

(a) $A = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 2 & -3 \\ 3 & 2 & 1 \\ 2 & -1 & -1 \end{bmatrix}$. *(b)* $A = \begin{bmatrix} 4 & -3 & 1 & 2 \\ 0 & -1 & 1 & 1 \\ 2 & -1 & 2 & 5 \\ 7 & -1 & 2 & 7 \end{bmatrix}$.

En los ejercicios 77 a 82 hallar los valores propios, vectores propios correspondientes y polinomio característico de la matriz A

77 $A = \begin{bmatrix} -1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 3 & -1 \end{bmatrix}$. **78** $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$. **79** $A = \begin{bmatrix} 2 & 1 \\ -2 & 4 \end{bmatrix}$. **80** $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & -1 \end{bmatrix}$.

81 $A = \begin{bmatrix} 4 & -3 & 1 & 2 \\ 0 & -1 & 1 & 1 \\ 2 & -1 & 2 & 5 \\ 7 & -1 & 2 & 7 \end{bmatrix}$ **82** $A = \begin{bmatrix} 8 & -3 & 3 & 9 \\ -3 & -2 & 0 & 0 \\ 3 & 0 & 4 & 7 \\ 9 & 0 & 7 & 14 \end{bmatrix}$

(a) Utilizar la función `rref` para encontrar la forma escalonada reducida equivalente a F y verificar que sus columnas son vectores linealmente independientes de $\mathbb{R}^4$.

(b) Hallar la factorización QR de la matriz F.

83 Sea $F = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$.

84 Utilizar la función `rref` para probar que los vectores $\vec{v}_1 = (1,-1,2,1)$, $\vec{v}_2 = (2,0,1,1)$ y $\vec{v}_3 = (1,1,-1,1)$ son linealmente independientes y encontrar una base ortonormal para el subespacio $\text{gn}(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ de $\mathbb{R}^4$.

Excel, la herramienta Solver y programación lineal (respuestas en página 1099)

En los ejercicios 85 a 106 utilizar la herramienta Solver para resolver el problema de programación lineal dado.

85 El problema del ejercicio 90 del capítulo 6.
86 El problema del ejercicio 92 del capítulo 6.
87 El problema del ejercicio 96 del capítulo 6.
88 El problema del ejercicio 100 del capítulo 6.
89 El problema del ejercicio 102 del capítulo 6.
90 El problema del ejercicio 104 del capítulo 6.
91 El problema del ejercicio 106 del capítulo 6.
92 El problema del ejercicio 107 del capítulo 6.
93 El problema del ejercicio 111 del capítulo 6.

94 Maximizar $P = 3x_1 - x_2 + 5x_3$

sujeto a
$$\begin{cases} x_1 + 2x_2 + x_3 \leq 5 \\ 4x_1 + 2x_2 - 5x_3 \geq 2 \\ 2x_1 + 4x_2 - x_3 \geq 4 \end{cases}$$
$x_1, x_2, x_3 \geq 0$

95 Minimizar $C = 2x_1 + 3x_2 + x_3$

sujeto a
$$\begin{cases} x_1 + 2x_2 + x_3 \leq 5 \\ 4x_1 + 2x_2 - 3x_3 = 7 \\ 5x_1 + 2x_2 - x_3 \geq 4 \end{cases}$$
$x_1, x_2, x_3 \geq 0$

96 Maximizar $P = 2x_1 + 3x_2 + x_3$

sujeto a
$$\begin{cases} x_1 + 2x_2 + x_3 \leq 5 \\ 4x_1 + 2x_2 - 3x_3 = 7 \\ 5x_1 + 2x_2 - x_3 \geq 4 \end{cases}$$
$x_1, x_2, x_3 \geq 0$

97 Maximizar $P = 3x_1 + 4x_2 + 2x_3 + 5x_4$

sujeto a
$$\begin{cases} 10x_1 + 3x_2 + 8x_3 + 2x_4 \geq 200 \\ 18x_1 + 10x_2 + 15x_3 + 15x_4 \geq 190 \\ 45x_1 + 25x_2 + 20x_3 + 37x_4 \geq 450 \\ x_1 + 3x_2 + 2x_3 + x_4 = 30 \end{cases}$$
$x_1, x_2, x_3, x_4 \geq 0$

98 Minimizar $C = 3x_1 + 4x_2 + 2x_3 + 5x_4$

sujeto a
$$\begin{cases} 10x_1 + 3x_2 + 8x_3 + 2x_4 \geq 200 \\ 18x_1 + 10x_2 + 15x_3 + 15x_4 \geq 190 \\ 45x_1 + 25x_2 + 20x_3 + 37x_4 \geq 450 \\ x_1 + 3x_2 + 2x_3 + x_4 = 30 \end{cases}$$
$x_1, x_2, x_3, x_4 \geq 0$

99 Maximizar $P = x_1 + 2x_2 + 3x_3 + 5x_4$

sujeto a
$$\begin{cases} x_1 + 3x_2 + x_3 + 2x_4 \geq 4 \\ 6x_1 + 5x_2 + 7x_3 + 8x_4 \geq 8 \\ 3x_1 + x_2 + 2x_3 + 2x_4 \geq 10 \\ 2x_1 + x_2 + 3x_3 + 2x_4 \leq 7 \end{cases}$$
$x_1, x_2, x_3, x_4 \geq 0$

100 Minimizar $C = x_1 + 2x_2 + 3x_3 + 5x_4$

sujeto a
$$\begin{cases} x_1 + 3x_2 + x_3 + 2x_4 \geq 4 \\ 6x_1 + 5x_2 + 7x_3 + 8x_4 \geq 8 \\ 3x_1 + x_2 + 2x_3 + 2x_4 \geq 10 \\ 2x_1 + x_2 + 3x_3 + 2x_4 \leq 7 \end{cases}$$
$x_1, x_2, x_3, x_4 \geq 0$

101 Minimizar $C = x_1 + x_2 + 4x_3 + 2x_4$

sujeto a
$$\begin{cases} x_1 + 4x_2 - x_3 + 2x_4 \leq 3 \\ 2x_1 - x_2 + x_3 + 8x_4 \leq 7 \\ 3x_1 + x_2 + 2x_3 + 2x_4 \leq 10 \\ 2x_1 + x_2 + 3x_3 + 2x_4 \geq 7 \\ x_1 + 2x_2 + x_3 + x_4 \leq 5 \end{cases}$$
$x_1, x_2, x_3, x_4 \geq 0$

102 Maximizar $P = x_1 + x_2 + 4x_3 + 2x_4$

sujeto a
$$\begin{cases} x_1 + 4x_2 - x_3 + 2x_4 \leq 3 \\ 2x_1 - x_2 + x_3 + 8x_4 \leq 7 \\ 3x_1 + x_2 + 2x_3 + 2x_4 \leq 10 \\ 2x_1 + x_2 + 3x_3 + 2x_4 \geq 7 \\ x_1 + 2x_2 + x_3 + x_4 \leq 5 \end{cases}$$
$x_1, x_2, x_3, x_4 \geq 0$

103 Maximizar $P = x_1 + 2x_2 + 3x_3 + 5x_4$

sujeto a
$$\begin{cases} x_1 + 3x_2 + x_3 + 2x_4 \geq 4 \\ 2x_1 - x_2 + x_3 + 8x_4 \leq 8 \\ 3x_1 + x_2 + 2x_3 + 2x_4 \leq 10 \\ 2x_1 + x_2 + 3x_3 + 2x_4 \leq 7 \\ x_1 + 2x_2 + x_3 + x_4 \leq 9 \end{cases}$$
$x_1, x_2, x_3, x_4 \geq 0$

104 Minimizar $C = x_1 + 2x_2 + 3x_3 + 5x_4$

sujeto a
$$\begin{cases} x_1 + 3x_2 + x_3 + 2x_4 \geq 4 \\ 2x_1 - x_2 + x_3 + 8x_4 \leq 8 \\ 3x_1 + x_2 + 2x_3 + 2x_4 \leq 10 \\ 2x_1 + x_2 + 3x_3 + 2x_4 \leq 7 \\ x_1 + 2x_2 + x_3 + x_4 \leq 9 \end{cases}$$
$x_1, x_2, x_3, x_4 \geq 0$

105 Maximizar $P = 3x_1 - 2x_2 + x_3 + x_4 + 2x_5$

sujeto a
$$\begin{cases} x_1 - x_2 + x_3 + 2x_4 + x_5 \leq 20 \\ 3x_1 + 2x_2 + x_3 - 2x_4 + 3x_5 \geq 10 \\ 5x_1 - x_2 + x_3 + 2x_4 + x_5 \geq 9 \\ 4x_1 + x_2 + x_3 + x_4 + x_5 \leq 8 \end{cases}$$
$x_1, x_2, x_3, x_4, x_5 \geq 0$

106 Minimizar $C = 3x_1 - 2x_2 + x_3 + x_4 + 2x_5$

sujeto a
$$\begin{cases} x_1 - x_2 + x_3 + 2x_4 + x_5 \leq 20 \\ 3x_1 + 2x_2 + x_3 - 2x_4 + 3x_5 \geq 10 \\ 5x_1 - x_2 + x_3 + 2x_4 + x_5 \geq 9 \\ 4x_1 + x_2 + x_3 + x_4 + x_5 \leq 8 \end{cases}$$
$x_1, x_2, x_3, x_4, x_5 \geq 0$

8 Álgebra lineal numérica

Una exposición del ánalisis lineal, como la que se ha hecho a lo largo de este texto, resultaría incompleta si no se muestran las dificultades inherentes que se presentan al realizar los cálculos numéricos en esta materia y la forma de resolverlas matemáticamente. Este capítulo contiene una introducción a los principales métodos del análisis numérico, la ciencia y el arte de calcular, que se utilizan en álgebra lineal. La primera sección trata brevemente el tema de aritmética de punto flotante en las computadoras y los tipos de errores que implica. En la segunda sección, se estudian los métodos directos para resolver sistemas lineales con el propósito de automatizar, y minimizar errores de redondeo. En la tercera sección se presentan algunos métodos iterativos para aproximar soluciones de sistemas lineales e inversas de matrices. La cuarta sección contiene el estudio de las transformaciones de Householder, que se utilizan para reducir matrices a un formato similar pero más simple. Y en la quinta sección veremos el apasionante tema del cálculo numérico de valores y vectores propios de matrices. En la última sección se incluye un conjunto de ejercicios propuestos al lector para su resolución.

8.1 Aritmética de la computadora y errores de redondeo

Los dispositivos electrónicos para hacer cálculos numéricos, computadoras y calculadoras, necesitan representar los números reales en determinado formato para poder operarlos aritméticamente. Además, independientemente de las respectivas capacidades, cada computadora está limitada a expresar sólo un número finito de números racionales; esto es, de números con expansión decimal finita. Por ende, en toda máquina se cometen cierto tipo de errores al hacer cómputo numérico. En esta sección daremos una breve introducción de la forma en la que estos dispositivos representan números reales mediante aritmética de punto flotante y a identificar los errores que eso conlleva.

Representación de números en aritmética de punto flotante

Las computadoras realizan la mayor parte del trabajo de cómputo en aritmética de **punto flotante**. Un número en este formato se escribe como[1]

$$\pm 0.\, d_1 d_2 \cdots d_n \times 10^r \tag{8.1}$$

[1] En realidad, la mayoría de las computadoras trabajan en sistema numérico binario, pero, por simplicidad en la exposición, hemos preferido utilizar el sistema decimal con el que todos estamos acostumbrados.

donde r es un número entero, con $-M_1 \leq r \leq M_2$; cada d_i es alguno de los dígitos 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 y $d_1 \neq 0$. Las cotas, M_1 y M_2, del **exponente** r, dependen de la computadora y del software que se utilicen; mientras que el entero n en (8.1) es el **número de cifras significativas**. A la sucesión $d_1 d_2 \cdots d_n$ en (8.1) se le llama **mantisa** del número en punto flotante. Por ejemplo, si $n = 4$, entonces

$$\begin{array}{lll} +0.1416 \times 10^2, & -0.1011 \times 10^{-5}, & +0.9321 \times 10^{-11}, \\ -0.3457 \times 10^0, & -0.1020 \times 10^2, & +0.3141 \times 10^1 \end{array}$$

son números en aritmética de punto flotante con $n = 4$ cifras significativas.

Ahora supongamos que se requiere representar un número R en aritmética de punto flotante con n cifras significativas. Para ello primero escribimos el número en la forma

$$R = \pm 0.d_1 d_2 \cdots d_n d_{n+1} \cdots \times 10^{-r} \tag{8.2}$$

donde los $d_i \in \{0, 1, 3, 4, 5, 6, 7, 8, 9\}$, $d_1 \neq 0$ y r es un entero que satisface $-M_1 \leq r \leq M_2$, donde M_1 y M_2 son las cotas que mencionamos antes. Puesto que el número debe tener exactamente n cifras significativas, es necesario limitar la expansión decimal, $0.d_1 d_2 \cdots d_n d_{n+1} \cdots$, a n dígitos representativos. Tenemos dos opciones:

Truncamiento. Se conservan las primeras n cifras decimales de (8.2). Entonces, con este criterio,

$$R = \pm 0.d_1 d_2 \cdots d_n \times 10^{-r}$$

Se dice, en este caso, que el número R en (8.2) se ha truncado a n cifras significativas.

Redondeo. Si $d_{n+1} \geq 5$, entonces la representación para R es

$$R = \pm(0.d_1 d_2 \cdots d_n + 10^{-n}) \times 10^{-r}$$

es decir, se redondea como se hace usualmente en educación elemental y se trunca el resultado a n cifras significativas. Si $d_{n+1} < 5$, entonces R se trunca a n cifras significativas. Se dice, en este caso, que el número R en (8.2) se ha redondeado a n cifras significativas

▶ **Ejemplo 8.1** Si $R = 8/3$, representar R en aritmética de punto flotante con cuatro cifras significativas *(a)* mediante truncamiento y *(b)* por redondeo.◀

Solución

$$\begin{aligned} R &= \tfrac{8}{3} \\ &= 2.6666666\cdots \\ &= +0.26666666\cdots \times 10^1 \end{aligned}$$

(a) Se trunca a 4 cifras significativas:

$$R = +0.2666 \times 10^1$$

(b) En este caso $n = 4$ y $d_{n+1} = d_5 = 6 \geq 5$, por tanto,

$$R = +0.2667 \times 10^1 \quad ✔$$

▶ **Ejemplo 8.2** Redondear los siguientes números a 4 cifras significativas:

1. 1.9983
2. 1.9985
3. 1.9996◀

Solución 1. La quinta cifra después del punto decimal es menor que 5, por tanto, 1.998 es la representación (aproximación) con cuatro cifras significativas para este caso.

2. La quinta cifra después del punto decimal es mayor que 5, por tanto, al sumar 1 a la cuarta cifra, se obtiene la representación 1.999.
3. La quinta cifra después del punto decimal es mayor que 5, por tanto, al sumar 1 a la cuarta cifra se obtiene

$$1.999 + 0.001 = 2$$

como representación (aproximación) en aritmética de punto flotante con cuatro cifras significativas para el número 1.9996 ✔

Errores de redondeo

Existen dos maneras de medir el error que se comete al aproximar una cantidad.

Definición 8.1 *Sea R^* una aproximación para R.*

1. *El **error absoluto** de aproximar R con R^* se define como*
$$|R - R^*|.$$
2. *El **error relativo de aproximar** R con R^* es*
$$\frac{|R - R^*|}{|R|}$$
si $R \neq 0$.

El error absoluto o relativo que se comete al aproximar un número R por su representación en punto flotante se llama **error de redondeo**, sin importar si se hizo por truncamiento o por redondeo.

Por lo general, el error absoluto es el que más se emplea para medir errores de aproximación; sin embargo, cuando la cantidad R es muy grande o muy pequeña, el error relativo es más significativo que el error absoluto. Por ejemplo, supongamos que la cantidad exacta es $R = 0.32 \times 10^{-5}$ y $R^* = 0.16 \times 10^{-5}$, entonces

$$|R - R^*| = 0.16 \times 10^{-5}$$

el cual es, en apariencia, muy pequeño. Pero, al medir el error relativo

$$\frac{|R - R^*|}{|R|} = 0.5$$

nos damos cuenta que en realidad el error absoluto que se comete es de 50 % de la cantidad R. De manera análoga, si $R = 0.345674345 \times 10^7$ y $R^* = 0.345600000 \times 10^7$, entonces el error absoluto

$$|R - R^*| = 0.74345 \times 10^3$$

es aparentemente muy grande; mientras que el error relativo

$$\frac{|R - R^*|}{|R|} = 2.1507 \times 10^{-4}$$

nos indica que en realidad el error absoluto es, aproximadamente, sólo 0.022 % de la cantidad R.

❍ **Nota 8.1** Algunos paquetes permiten, o tienen por defecto, el uso de doble precisión en los cálculos; esto es, trabajan con el doble de cifras significativas que en la precisión simple. MATLAB, por ejemplo, es un paquete que trabaja por defecto con doble precisión.

8.2 Métodos directos para resolver sistemas lineales

En el capítulo 1 de este libro vimos el método de Gauss para resolver sistemas lineales de ecuaciones. Recordemos que para encontrar las soluciones del sistema $A_1\vec{x} = \vec{b}$, se forma la matriz ampliada del sistema $A = [A_1\ \vec{b}]$, se lleva a una forma escalonada y se aplica sustitución regresiva al sistema escalonado que se obtiene. En los capítulos precedentes hemos resuelto sistemáticamente sistemas lineales utilizando este método, aplicando estrategias particulares y eligiendo las operaciones de renglón y pivotes como ha sido conveniente en cada problema; sin embargo, para sistemas grandes, se requiere de la automatización del método en algún algoritmo que se pueda programar para que lo ejecute una computadora y que se aplique, al menos en teoría, a cualquier sistema lineal de ecuaciones con la mínima intervención de inteligencia humana. Nos limitaremos en lo que resta de este capítulo a sistemas lineales cuadrados; esto es, a sistemas lineales de n ecuaciones y n incógnitas.

8.2.1 Método de Gauss para sistemas lineales de orden n con sustitución regresiva

Sea A_1 una matriz cuadrada de orden n y consideremos el sistema lineal de ecuaciones

$$A_1\vec{x} = \vec{b} \tag{8.3}$$

Existen tres posibilidades:

1. El sistema tiene solución única.
2. El sistema no tiene solución.
3. El sistema tiene una infinidad de soluciones.

La alternativa del inciso 1 se tiene cuando la matriz A_1 es equivalente a la identidad; mientras que las alternativas de los incisos 2 o 3 suceden cuando la matriz A_1 no es equivalente a la identidad (cfr. teo-

rema 1.4, pág. 24, el tercer criterio de la página 31 o el teorema 2.3 de la pág. 67). Por esta razón, cuando se afirma que el sistema (8.3) no tiene solución única, significa que la alternativa del inciso 2 o del 3 son verdaderas. En el siguiente algoritmo solamente trataremos sistemas con solución única; indicando que se da uno de los casos 2 o 3 al señalar simplemente que *el sistema no tiene solución única*.

Método de solución de sistemas lineales por eliminación gaussiana con sustitución regresiva

Algoritmo 8.1

Entrada: *La matriz aumentada* $A = [A_1\ \vec{b}]$ *del sistema (8.3), donde* $A_1 = [a_{ij}]$ *es la matriz de coeficientes y* $\vec{b} = [a_{1,n+1}\ a_{2,n+1}\ \cdots\ a_{n,\,n+1}]^t$ *es el término independiente.*
Salida: *La solución* $x_1\ x_2\ \cdots\ x_n$ *del sistema (8.3) o la leyenda "El sistema no tiene solución única".*

Para cada $j = 1, \ldots, n-1$*, de manera recurrente, se genera una matriz* $A^{(j)} = [a^{(j)}_{\alpha\beta}]$ *equivalente por filas a la matriz A, donde la notación* $a^{(0)}_{\alpha\beta}$ *representa la componente* $a_{\alpha\beta}$ *de A, de la siguiente manera (pasos 1 a 4):*

Paso 1. *Se elige el menor subíndice* $k \geq j$ *de la columna* j *tal que* $a^{(j-1)}_{kj} \neq 0$.

Paso 2. *Si no existe k; es decir, si* $a^{(j-1)}_{ij} = 0$ *para todo* $i \geq j$*, se interrumpe la generación de la matriz* $A^{(j)}$ *y se despliega la salida "El sistema no tiene solución única".*

Paso 3. *Si k existe, el proceso continúa. Se calculan*

$$m^{(j)}_{ij} = \frac{a^{(j-1)}_{ij}}{a^{(j-1)}_{kj}} \tag{8.4}$$

$i = k+1, \ldots, n$.

Paso 4. *Mediante la operación de renglón*

$$R_i \leftrightarrow R_i - m^{(j)}_{ij} R_k$$

se transforman en ceros todos los elementos por debajo de la componente $a^{(j-1)}_{kj}$ *en la matriz* $A^{(j-1)}$*; a la matriz resultante se le aplica la operación de renglón* $R_j \leftrightarrow R_k$ *para obtener la matriz* $A^{(j)}$.

Paso 5. *Si el proceso continúa sin interrupción hasta* $j = n-1$*, y k existe para* $j = n$*, es decir,* $a^{(n-1)}_{nn} \neq 0$*, sean*

$$x_n = \frac{a^{(n-1)}_{n,\,n+1}}{a^{(n-1)}_{nn}}$$

y

$$x_i = \frac{1}{a^{(n-1)}_{ii}} \left[a^{(n-1)}_{i,\,n+1} - \sum_{j=i+1}^{n} a^{(n-1)}_{ij} x_j \right]$$

para $i = n-1, \ldots, 1$*. Entonces la salida es la matriz fila*

$$[x_1\ x_2\ \cdots\ x_n]$$

► **Ejemplo 8.3** Resolver, mediante el método de solución de sistemas lineales por eliminación gaussiana con sustitución regresiva, el sistema

$$\begin{array}{rcrcrcr} & & 2x_2 & - & x_3 & = & -1 \\ x_1 & - & x_2 & + & 3x_3 & = & 4 \\ 2x_1 & + & 3x_2 & - & x_3 & = & 6 \end{array}$$ ◄

Solución En este caso

$$A^{(0)} = A = \begin{bmatrix} 0 & 2 & -1 & -1 \\ 1 & -1 & 3 & 4 \\ 2 & 3 & -1 & 6 \end{bmatrix}.$$

$j=1$: el primer elemento no nulo de la primera columna de A es $a_{21}=1$, así que $k=2$; entonces $a_{k1}=a_{21}=1$ y $m_{31}=\frac{a_{31}}{a_{21}}=2$. Al hacer la operación de renglón $R_3 \leftrightarrow R_3 - m_{31}R_2$ a la matriz $A=A^{(0)}$ se obtiene la equivalencia

$$\begin{bmatrix} 0 & 2 & -1 & -1 \\ 1 & -1 & 3 & 4 \\ 2 & 3 & -1 & 6 \end{bmatrix} \sim \begin{bmatrix} 0 & 2 & -1 & -1 \\ 1 & -1 & 3 & 4 \\ 0 & 5 & -7 & -2 \end{bmatrix};$$

y al aplicar el intercambio de filas $R_1 \leftrightarrow R_2$ a la última matriz resulta

$$A^{(1)} = \begin{bmatrix} 1 & -1 & 3 & 4 \\ 0 & 2 & -1 & -1 \\ 0 & 5 & -7 & -2 \end{bmatrix}.$$

$j=2$: el primer elemento no nulo de la columna 2 —por debajo de a_{22} o igual a él— de la matriz $A^{(1)}$ es $a_{22}=2$, así $k=2$; entonces $a_{k2}=a_{22}=2$ y $m_{32}=\frac{a_{32}}{a_{22}}=\frac{5}{2}$. Al hacer la operación de renglón $R_3 \leftrightarrow R_3 - m_{32}R_2$ a la matriz $A^{(1)}$ se obtiene la equivalencia

$$\begin{bmatrix} 1 & -1 & 3 & 4 \\ 0 & 2 & -1 & -1 \\ 0 & 5 & -7 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 3 & 4 \\ 0 & 2 & -1 & -1 \\ 0 & 0 & -9/2 & 1/2 \end{bmatrix};$$

y al aplicar el intercambio de filas $R_2 \leftrightarrow R_2$ a la última matriz resulta (obviamente esta operación en la práctica, para el presente caso, es innecesaria):

$$A^{(2)} = \begin{bmatrix} 1 & -1 & 3 & 4 \\ 0 & 2 & -1 & -1 \\ 0 & 0 & -9/2 & 1/2 \end{bmatrix}.$$

Puesto que k existe para $j=3$ ($a_{33}^{(2)}=-9/2\neq 0$), se tiene

$$x_3 = \frac{a_{34}^{(2)}}{a_{33}^{(2)}} = \frac{1/2}{-9/2} = -\frac{1}{9};$$

$$x_2 = \frac{1}{a_{22}^{(2)}}\left[a_{24}^{(2)} - \sum_{j=2+1}^{3} a_{2j}^{(n-1)}x_j\right] = \frac{1}{a_{22}^{(2)}}\left[a_{24}^{(2)} - a_{23}^{(2)}x_3\right] = \frac{1}{2}[-1-(-1)(-1/9)] = -\frac{5}{9};$$

$$
\begin{aligned}
x_1 &= \frac{1}{a_{11}^{(2)}}\left[a_{14}^{(2)} - \sum_{j=1+1}^{3} a_{1j}^{(2)} x_j\right] \\
&= \frac{1}{1}(4-((-1)(-5/9)+3(-1/9)) = \frac{34}{9}\,.
\end{aligned}
$$

Por tanto, la salida es

$$[x_1\; x_2\; x_3] = [34/9\;\; -5/9\;\; -1/9]. \quad ✔$$

▶ **Ejemplo 8.4** Resolver, mediante el método de solución de sistemas lineales por eliminación gaussiana con sustitución regresiva, el sistema

$$
\begin{array}{rcrcrcr}
x_1 & - & x_2 & + & 3x_3 & = & -1 \\
-x_1 & + & x_2 & + & 2x_3 & = & 4 \\
3x_1 & - & 3x_2 & + & 2x_3 & = & -2\,.
\end{array} ◀
$$

Solución En este caso

$$
A = \begin{bmatrix} 1 & -1 & 3 & -1 \\ -1 & 1 & 2 & 4 \\ 3 & -3 & 2 & -2 \end{bmatrix}.
$$

$j = 1$: $k = 1$, $a_{11}^{(1)} = 1$, y

$$
\begin{bmatrix} 1 & -1 & 3 & -1 \\ -1 & 1 & 2 & 4 \\ 3 & -3 & 2 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 3 & -1 \\ 0 & 0 & 5 & 3 \\ 0 & 0 & -7 & 1 \end{bmatrix} = A^{(2)}.
$$

$j = 2$: Para este índice no existe k; es decir, $a_{i2}^{(2)} = 0$, $i = 2, 3$. Por tanto, el proceso debe detenerse y *el sistema no tiene solución única*. ✔

Programa en MATLAB para el método de Gauss

En esta parte implementaremos, en un programa en MATLAB que llamaremos `sistemas_gauss`, el método de solución de sistemas lineales por eliminación gaussiana con sustitución regresiva. Para ello es conveniente que el lector interesado en este tema repase (o estudie), si así se requiere, el material de la sección 7.2; en particular, el programa hará uso de las funciones `intercambiofilas` y `sumafilas` programadas en el apartado 7.2.6.

La parte del programa `sistemas_gauss` que determina el primer elemento no nulo de una columna, la llevaremos a cabo mediante la función `primera_entrada_nonula` contenida en la figura 8-1. Las instrucciones para el uso de esta función vienen dadas en los comentarios al inicio del guión. En MATLAB si una matriz `X` se compara con una matriz `Y`, mediante algunos de los comandos de comparación lógica `>=`, `<=`, `==`, etc., entonces crea una matriz que tiene unos donde la comparación es verdadera y ceros donde es falsa; así, por ejemplo, si `X=[1 -1 3]`, `Y=[2 -2 3]`, la salida que produce MATLAB a la instrucción `Z=(X<=Y)` es `Z=[1 0 1]`. Por otra parte, si a las operaciones aritméticas en MATLAB, `*`, `/`, `\`, `^`, les precede un punto, entonces se realizan elemento a elemento en los arreglos que operan; por ejemplo, si `X`, `Y` son los vectores de arriba, entonces la operación `X.*Y` produce la matriz `[2 2 9]`; mientras que la instrucción `X.^Y` tiene la salida `[1 1 27]`. La función reservada en

```
function z=primera_entrada_nonula(X,p)
%Esta función da como salida la primera
%coordenada z=X_k, k>=p, distinta de cero,
%de la matriz fila X=[X_1 X_2 ... X_n]
%si es que existe; en caso contrario
%la salida será z=0.
n=length(X);
m=1:n;
X=X.*(m>=p);
if X==zeros(1,n)
   z=0;
 else
   u=(n+1)*(X==0);
   v=m.*(X~=0);
   z=min(u+v);
end;
```

Figura 8-1 • Función primera_entrada_nonula.

MATLAB, min, da como resultado el valor mínimo de la matriz fila o columna donde se evalúe. Con esta información, el lector debe intentar comprender por sí mismo las instrucciones* de la figura 8-1 para la programación de esta función y verificar en MATLAB su correcto funcionamiento. La última parte del programa sistemas_gauss, se realizará con la función sust_regresiva, cuya programación está contenida en la figura 8-2; las instrucciones para su uso vienen dadas al inicio del guión. El comando sum calcula la suma de las componentes de la matriz fila o columna donde se evalúe; la instrucción B(:,p) = [] elimina la columna p de la matriz B. Nuevamente, con la precedente información, el lector debe entender por sí mismo las instrucciones del programa de la figura 8-2 en MATLAB para verificar su funcionamiento. Finalmente, las instrucciones del programa sistemas_gauss están escritas en el guión contenido en la figura 8-3. Las indicaciones de su uso están dadas al inicio del

```
function s=sust_regresiva(M)
%***sust_regresiva(M)***
%Esta función resuelve, por sustitución
%regresiva, sistemas cuadrados con matriz
%ampliada M=[N b], donde N es triangular superior
%y los elementos de su diagonal son no nulos.
n=size(M,1);
m=1:n;
x=zeros(1,n+1);
x(n)=M(n,n+1)/M(n,n);
for p=1:n-1
   a=1/M(n-p,n-p);
   suma=sum((M(n-p,:).*([m~=n-p 0])).*x);
   x(n-p)=a*(M(n-p,n+1)-suma);
end
x(:,n+1)=[];
s=x;
```

Figura 8-2 • Función sust_regresiva.

```
function M=sistemas_gauss(A)
%***sistemas_gauss(A)***
%Este programa resuelve, por el método de
%Gauss, el sistema cuadrado
%Bx=b, donde A=[B b] es la matriz ampliada
%del sistema. El programa emite un mensaje
%cuando el sistema tiene una infinidad
%de soluciones o es inconsistente y da
%como resultado la solución:
%    x1 x2 ... xn
%cuando es única.
%Utiliza las funciones primera_entrada_nonula,
%sust_regresiva, intercambiofilas y sumafilas
n=size(A,1);
A1=A;
for j=1:n
X=A1(:,j)';
k=primera_entrada_nonula(X,j);
if k~=0
   for i=k+1:n
   m=-A1(i,j)/X(k);
   A1=sumafilas(A1,i,k,1,m);
   end
   A1=intercambiofilas(A1,k,j);
   M1=A1;
else
   M1=0;
   break
end
end
if M1==0

   M= 'El sistema no tiene solución única';
else
   M=sust_regresiva(M1);
end
```

Figura 8-3 • Función sistemas_gauss.

* N. del autor: Por cuestiones de espacio, los programas se escribieron en dos columnas, pero las instrucciones en la segunda columna en realidad deben ser escritas, línea tras línea, después de la última instrucción de la primera columna, para su correcto funcionamiento en MATLAB.

guión. Recordamos al lector que todos los guiones de estas funciones deben ser escritos en el editor de MATLAB (de preferencia) y guardar todos ellos con el mismo nombre de la función y con extensión `.m` en el directorio donde se vayan a usar (cfr. apartado 7.2.6).

▶ **Ejemplo 8.5** Resolver en MATLAB, utilizando la función `sistemas_gauss`, el sistema lineal

$$\begin{array}{rcrcrcrcr} & & x_2 & - & x_3 & + & x_4 & = & 5 \\ 2x_1 & - & x_2 & + & 4x_3 & - & x_4 & = & -16 \\ 4x_1 & & & + & 3x_3 & - & 2x_4 & = & -9 \\ -3x_1 & - & 2x_2 & - & 5x_3 & + & 2x_4 & = & 6 \end{array}$$ ◀

Solución

```
EDU» A=[0 1 -1 1 5; 2 -1 4 -1 -16;4 0 3 -2 -9;-3 -2 -5 2 6]
A =
     0     1    -1     1     5
     2    -1     4    -1   -16
     4     0     3    -2    -9
    -3    -2    -5     2     6
EDU» sistemas_gauss(A)
ans =
    -1     4    -3    -2
EDU»
```
✔

▶ **Ejemplo 8.6** Resolver en MATLAB, utilizando la función `sistemas_gauss`, el sistema lineal

$$\begin{array}{rcrcrcrcr} x_1 & - & 3x_2 & - & x_3 & - & 2x_4 & = & -1 \\ -x_1 & - & 2x_2 & - & 4x_3 & - & 3x_4 & = & 2 \\ 2x_1 & + & x_2 & + & 5x_3 & + & x_4 & = & -1 \\ 3x_1 & + & 2x_2 & + & 8x_3 & - & 2x_4 & = & 3 \end{array}$$ ✔

Solución

```
EDU» A=[1 -3 -1 -2 -1;-1 -2 -4 -3 2;2 1 5 1 -1;3 2 8 -2 3]
A =
     1    -3    -1    -2    -1
    -1    -2    -4    -3     2
     2     1     5     1    -1
     3     2     8    -2     3
EDU» sistemas_gauss(A)
ans =
El sistema no tiene solución única
EDU»
```
✔

8.2.2 Método de Gauss para hallar la inversa de una matriz

Es evidente que el algoritmo 8.1 requiere de unos pequeños cambios para que calcule la inversa de una matriz o determine si es singular (no invertible); pues únicamente hay que modificarlo para que resuelva los sistemas $A\vec{x} = \vec{e}_i$, para cada $i = 1, 2, \ldots, n$, donde $\vec{e}_i$ es la columna i de la matriz identidad I_n. El siguiente algoritmo contiene los ajustes necesarios para ello.

Algoritmo 8.2

Entrada: *La matriz cuadrada* $A = [a_{ij}]$ *de orden n cuya inversa se va a calcular.*
Salida: *La matriz inversa de A o la leyenda* "La matriz no es invertible".
Paso 1. *Se aplica el algoritmo 8.1 de los pasos 1 a 4 a la matriz aumentada* $[A \;\; I_n]$ *donde* I_n *es la matriz identidad de orden n, para obtener en cada etapa del proceso una matriz* $[A^{(\nu)} \;\; B^{(\nu)}]$ *equivalente por filas a la matriz* $[A \;\; I_n]$.
Paso 2. *Si en el proceso una de las columnas de la matriz resultante* $[A^{(\nu)} \;\; B^{(\nu)}]$ *tiene ceros en todos los elementos de la diagonal, inclusive hacia abajo, entonces se despliega el mensaje* "La matriz no es invertible".
Paso 3. *En caso contrario se obtendrá una matriz* $[A^{(m)} \;\; B^{(m)}]$ *con la matriz* $A^{(m)}$ *triangular superior y todos los elementos en la diagonal distintos de cero. Se aplica la sustitución regresiva, del paso 5 del algoritmo 8.1, a cada uno de los sistemas que se obtienen con la matriz del lado izquierdo de la ampliación con cada una de las columnas de la matriz* $B^{(m)}$*. La salida es la matriz que tiene por columnas cada una de las soluciones en forma sucesiva.*

Programa en MATLAB para calcular la inversa de una matriz con el método de Gauss con sustitución regresiva

En la figura 8-4 se encuentra el código de la función `Metodo_Gauss_Bis` que contiene unas ligeras adecuaciones al programa `sistemas_Gauss` para poder utilizar éste en forma interna en la función `Inversa_Gauss`; mientras que el código de esta función se encuentra en la figura 8-5.

```
function M=Metodo_Gauss_Bis(A)
n=size(A,1);
A1=A;
for j=1:n
X=A1(:,j)';
k=primera_entrada_nonula(X,j);
if k~=0
   for i=k+1:n
   m=-A1(i,j)/X(k);
   A1=sumafilas(A1,i,k,1,m);
   end
   A1=intercambiofilas(A1,k,j);
   M1=A1;
else
   M1=0;
   break
end
end
if M1==0
    M=0;
else
   M=sust_regresiva(M1);
end
```

Figura 8-4 • Función `Metodo_Gauss_Bis`.

```
function B=Inversa_Gauss(A)
%***Inversa_Gauss(A)***
%calcula la inversa de la matriz A
%con el algoritmo de Gauss
n=size(A,1);
I=eye(n);
for j=1:n
  z(:,j)=Metodo_Gauss_Bis([A I(:,j)])';
end
  if z==0
     B='La matriz no es invertible.';
  else
     B=z;
end
```

Figura 8-5 • Función `Inversa_Gauss`.

► **Ejemplo 8.7** Hallar, si existe, la inversa de la matriz $A = \begin{bmatrix} 1 & -2 & 2 \\ 1 & -1 & 1 \\ -1 & 2 & -1 \end{bmatrix}$ utilizando la función `Inversa_Gauss` en MATLAB.◄

Solución

```
EDU» A=[1 -2 2 1 -1 1 -1 2 -1]
A =
     1    -2     2
     1    -1     1
    -1     2    -1
EDU» Inversa_Gauss(A)
ans =
    -1     2     0
     0     1     1
     1     0     1
EDU»
```
✔

▶ **Ejemplo 8.8** Hallar, si existe, la inversa de la matriz

$$G=\begin{bmatrix} 1 & -2 & 2 \\ 1 & -1 & 1 \\ -1 & 2 & -2 \end{bmatrix}$$

utilizando la función `Inversa_Gauss` en MATLAB.◀

Solución

```
EDU» G=[1 -2 2;1 -1 1;-1 2 -2];
EDU» Inversa_Gauss(G)
ans =
La matriz no es invertible.
EDU»
```
✔

8.2.3 *Factorización LU*

Recordemos que una matriz cuadrada de orden m es elemental, si se obtiene de la matriz identidad I_m al aplicarle una sola operación de renglón (cfr. definición 2.3). En el teorema 2.4 del apartado 2.1.4 (pág. 71) vimos el efecto que tienen las matrices elementales cuando multiplican por la izquierda a una matriz; esto es: si a una matriz A de tamaño $m \times n$ se le aplica una operación de renglón, la matriz resultante es el producto EA, donde E es la matriz elemental $m \times m$ que se obtiene de la identidad al aplicarle la misma operación de renglón. También vimos (cfr. teorema 2.5 y discusión que le precede) que toda matriz elemental es una matriz invertible y su inversa es una matriz elemental y, en particular —que es el caso que aquí nos interesa— si la matriz elemental E se obtuvo de la identidad al aplicarle la operación de renglón $R_i \leftrightarrow R_i + \beta R_j$, entonces E^{-1} es la matriz elemental que se obtiene de la identidad al aplicarle la operación de renglón $R_i \leftrightarrow R_i - \beta R_j$.

Supongamos que A es una matriz cuadrada de orden n con la propiedad de que al aplicarle el método de Gauss, algoritmo 8.1, al sistema $A\vec{x} = \vec{b}$, no se hace ningún intercambio de filas. Representemos, como es natural, el lado derecho de la ampliación $[A^{(j)}\ \vec{b}]$ por $A^{(j)}$, $j = 1, 2, \ldots n-1$, en cada uno de los pasos del algoritmo 8.1. Así, en la primera etapa para obtener $A^{(1)}$, se calcula $m_{21}^{(1)} = a_{21}/a_{ii}$ y se aplica la operación de renglón $R_2 \leftrightarrow R_2 - m_{21}^{(1)} R_1$ para obtener la equivalencia de matrices $A \leftrightarrow A_2^{(1)}$; pero, por lo que se acaba de explicar, $A_2^{(1)} = E_2^{(1)} A$ donde $E_2^{(1)}$ es la matriz elemental

$$E_2^{(1)} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -m_{21}^{(1)} & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \end{bmatrix};$$

después se calcula $m_{31}^{(1)} = a_{31}/a_{11}$ y se aplica operación de renglón $R_3 \leftrightarrow R_3 - m_{31}^{(1)} R_1$ para obtener la equivalencia de matrices $A \leftrightarrow A_3^{(1)}$; pero, por lo que mencionamos arriba, $A_3^{(1)} = E_3^{(1)} A_2^{(1)}$ donde $E_3^{(1)}$ es la matriz elemental

$$E_3^{(1)} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ -m_{31}^{(1)} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \end{bmatrix}.$$

Y, por tanto,

$$\begin{aligned} A_3^{(1)} &= E_3^{(1)} A_2^{(1)} \\ &= (E_3^{(1)} E_2^{(1)}) A. \end{aligned}$$

Es fácil ver que

$$(E_3^{(1)} E_2^{(1)}) = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -m_{21}^{(1)} & 1 & 0 & 0 & \cdots & 0 \\ -m_{31}^{(1)} & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Continuando con el proceso, se obtiene finalmente la matriz

$$\begin{aligned} A^{(1)} &= (E_n^{(1)} \cdots E_3^{(1)} E_2^{(1)}) A \\ &= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -m_{21}^{(1)} & 1 & 0 & \cdots & 0 \\ -m_{31}^{(1)} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -m_{n1}^{(1)} & 0 & 0 & \cdots & 1 \end{bmatrix} A. \end{aligned}$$

Denotemos simplemente por E_1 la matriz del lado derecho de la precedente igualdad; esto es

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -m_{21}^{(1)} & 1 & 0 & \cdots & 0 \\ -m_{31}^{(1)} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -m_{n1}^{(1)} & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Entonces,

$$A^{(1)} = E_1 A.$$

De manera análoga tenemos $A^{(2)} = E_2 A^{(1)}$ donde

$$E_2 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & -m_{32}^{(2)} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -m_{n2}^{(2)} & 0 & & 1 \end{bmatrix}$$

y, por tanto,

$$A^{(2)} = (E_2 E_1) A.$$

Así, de forma inductiva, si para cada $j = k+1, \ldots, n$, $E_j^{(k)}$ es la matriz elemental

$$E_j^{(k)} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \vdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & \vdots & 0 & \cdots & \vdots & 0 \\ \vdots & \vdots & \vdots & 1 & 0 & \vdots & 0 & 0 \\ 0 & 0 & \cdots & -m_{j,k}^{(k)} & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

y

$$E_k = E_n^{(k)} E_{n-1}^{(k)} \cdots E_{k+1}^{(k)}$$

entonces,[2]

$$E_k = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \vdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & -m_{(k+1)k}^{(k)} & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & -m_{(k+2)k}^{(k)} & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & 1 & 0 \\ 0 & 0 & \cdots & -m_{nk}^{(k)} & 0 & \cdots & 0 & 1 \end{bmatrix} \quad (8.5)$$

[2] Es fácil probar, lo cual se deja de ejercicio al lector, que efectivamente el producto $E_n^{(k)} E_{n-1}^{(k)} \cdots E_{k+1}^{(k)}$ está dado por el lado derecho de 8.5.

con los $m_{jk}^{(k)}$ dados por (8.4) del algoritmo 8.1, y

$$\begin{aligned} A^{(k)} &= E_k A^{(k-1)} \\ &= (E_k E_{k-1} \cdots E_1)A. \end{aligned}$$

Por tanto, $A^{(n-1)} = (E_{n-1}E_{n-2}\cdots E_1)A$. Donde $A^{(n-1)}$ es la matriz triangular superior que se obtiene con el método de Gauss del algoritmo 8.1; esto es,

$$A^{(n-1)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{nn} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n-1)} \end{bmatrix} \tag{8.6}$$

Denotemos la matriz triangular superior por U; es decir $U = A^{(n-1)}$. Con esta notación

$$U = (E_{n-1}E_{n-2}\cdots E_1)A$$

Ahora bien, toda matriz elemental es invertible, y ya que E_k es producto de matrices elementales, entonces es invertible; puesto que el producto de matrices invertibles es una matriz no singular, podemos despejar la matriz de la precedente igualdad obteniendo

$$A = (E_1^{-1}E_2^{-1}\cdots E_{n-1}^{-1})U.$$

Ya que

$$\left(E_j^{(k)}\right)^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \vdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & \cdots & \vdots & 0 \\ \vdots & \vdots & \vdots & 1 & 0 & \vdots & 0 & 0 \\ 0 & 0 & \cdots & -m_{jk}^{(k)} & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \vdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & \cdots & \vdots & 0 \\ \vdots & \vdots & \vdots & 1 & 0 & \vdots & 0 & 0 \\ 0 & 0 & \cdots & m_{jk}^{(k)} & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

para cada $j = k+1, \ldots, n$, se tiene

$$E_j^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \vdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & \cdots & \vdots & 0 \\ \vdots & \vdots & \vdots & 1 & 0 & \vdots & 0 & 0 \\ 0 & 0 & \cdots & m_{j,k} & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

y, por ende,

$$E_k^{-1} = (E_{k+1}^{(k)})^{-1} \cdots (E_{n-1}^{(k)})^{-1} (E_n^{(k)})^{-1}$$
$$= \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \vdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & \cdots & \vdots & 0 \\ \vdots & \vdots & \vdots & 1 & 0 & \vdots & 0 & 0 \\ 0 & 0 & \cdots & m_{(k+1)k}^{(k)} & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & m_{(k+2)k}^{(k)} & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & m_{nk}^{(k)} & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

No es difícil probar que si $L = (E_1^{-1} E_2^{-1} \cdots E_{n-1}^{-1})$, entonces

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ m_{21}^{(1)} & 1 & \vdots & 0 & 0 & \cdots & 0 & 0 \\ m_{31}^{(1)} & m_{32}^{(2)} & \ddots & 0 & 0 & \cdots & \vdots & 0 \\ \vdots & \vdots & \vdots & 1 & 0 & \vdots & 0 & 0 \\ m_{(k+1)1}^{(1)} & m_{(k+1)2}^{(2)} & \cdots & m_{(k+1)k}^{(k)} & 1 & 0 & \cdots & 0 \\ m_{(k+2)1}^{(1)} & m_{(k+2)2}^{(k)} & \cdots & m_{(k+2)k}^{(k)} & m_{(k+2)(k+1)}^{(k+1)} & 1 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n1}^{(1)} & m_{n2}^{(2)} & \cdots & m_{nk}^{(k)} & m_{n(k+1)}^{(k+1)} & \cdots & m_{(n-1)(n-1)}^{(n-1)} & 1 \end{bmatrix} \quad (8.7)$$

De esta manera hemos podido factorizar la matriz A como el producto de una matriz triangular inferior y una matriz triangular superior:

$$A = LU.$$

Hemos probado así el teorema que a continuación enunciamos.

Teorema 8.1 *Sea A una matriz cuadrada de orden n. Se supone que en todo el proceso del método de Gauss del algoritmo 8.1 los valores en la diagonal de las matrices* $A^{(k)}$ *son distintos de cero. Entonces A se puede factorizar en la forma*

$$A = LU \tag{8.8}$$

donde la matriz U es triangular superior y la matriz L es triangular inferior. De hecho, U está dada por el lado derecho de la igualdad (8.6) y L por (8.7).

Cuando se aplique el teorema precedente, diremos que (8.8) es la factorización LU de la matriz A.

▶ **Ejemplo 8.9** Encontrar la factorización LU de la matriz

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & -1 & 2 \\ 3 & -1 & 5 \end{bmatrix}. ◀$$

Solución Apliquemos el algoritmo 8.1 para llevar la matriz a una forma escalonada equivalente:

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & -1 & 2 \\ 3 & -1 & 5 \end{bmatrix} \underset{\substack{R_2 \leftrightarrow R_2 - 2R_1 \\ R_3 \leftrightarrow R_3 - 3R_1}}{\longleftrightarrow} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 2 & -1 \end{bmatrix}$$
$$\underset{R_2 \leftrightarrow R_3 - 2R_2}{\longleftrightarrow} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

De donde $m_{21}^{(1)} = 2$, $m_{31}^{(1)} = 3$ y $m_{32}^{(2)} = 2$. Por tanto,

$$U = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

y

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}.$$

Comprobación:

$$\begin{aligned} LU &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 & 2 \\ 2 & -1 & 2 \\ 3 & -1 & 5 \end{bmatrix} \\ &= A. \quad ✔ \end{aligned}$$

Factorización LU y sistemas lineales

La factorización LU es muy útil cuando se resuelven sistemas lineales con la misma matriz de coeficientes y distintos términos independientes; pues la sustitución regresiva y la sustitución hacia adelante son muy eficientes para matrices que están en forma triangular superior e inferior, respectivamente. Observemos que para resolver un sistema lineal con un solo término independiente, no tiene caso utilizar factorización LU; ya que para ello se requiere aplicar el método de Gauss para encontrar la matriz U que es básicamente el mismo esfuerzo que resolver el sistema. Ilustraremos cómo se resuelven sistemas lineales con un solo término independiente utilizando esta factorización; teniendo en mente que la utilidad del método radica en que se aplica en forma simultánea a sistemas con distintos términos independientes pero con la misma matriz de coeficientes.

Supongamos que para la matriz A ya se tiene la factorización LU. Entonces

$$A = LU$$

y, por tanto,

$$A\vec{x} = \vec{b}$$

equivale a

$$L(U\vec{x}) = \vec{b}.$$

Si $\vec{y} = U\vec{x}$, se obtiene el sistema

$$L\vec{y} = \vec{b}$$

con la matriz L triangular inferior; se resuelve este sistema, por sustitución hacia adelante y la solución se sustituye en el sistema $\vec{y} = U\vec{x}$ que es triangular superior y se hace sustitución regresiva.

▶ **Ejemplo 8.10** Resolver el sistema

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & -1 & 2 \\ 3 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix}$$

utilizando la factorización LU de la matriz A del ejemplo 8.9.◀

Solución La matriz de coeficientes del sistema es la matriz A del ejemplo 8.9, con factorización

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 3 \end{bmatrix}}_{U}$$

Resolvamos el sistema $L\vec{y} = \vec{b}$; es decir, el sistema

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix},$$

entonces

$$\begin{aligned} y_1 &= -1, \\ y_2 &= -2-2y_1 \\ &= -2-2(-1) \\ &= 0, \\ y_3 &= -3y_1-2y_2 \\ &= -3(-1)-2(0) \\ &= 3. \end{aligned}$$

Ahora resolvamos el sistema $U\vec{x}=\vec{y}$; esto es, el sistema

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix},$$

entonces

$$\begin{aligned} x_3 &= 1, \\ x_2 &= 2x_3 \\ &= 2, \\ x_1 &= -1+x_2-2x_3 \\ &= -1+2-2(1) \\ &= -1. \quad ✔ \end{aligned}$$

Factorización LU y matriz de permutación

No siempre es posible aplicar el método de Gauss sin tener que intercambiar filas en una matriz A; aun así, se puede encontrar una factorización LU con un factor matricial adicional, una **matriz de permutación** P; esto es

$$PA = LU$$

o

$$\begin{aligned} A &= P^{-1}LU \\ &= P^{t}LU \end{aligned}$$

donde L es una matriz triangular inferior, U es una matriz triangular superior y P es una matriz que tiene “registrados” los intercambios de renglón necesarios —de ahí su nombre—. Supongamos que en el proceso de aplicar el algoritmo 8.1 a la matriz A, para llevar ésta a una forma escalonada equivalente, se tienen que hacer intercambios de renglones; entonces hay un número finito de matrices elementales $M_1, M_2, \ldots, M_k$ —cada una se obtiene intercambiando los mismos renglones a la matriz identidad en forma simultánea a los intercambios que se realizan en el proceso a la matriz A— tales que

$$A = M_kM_{k-1}\cdots M_1.$$

De esta manera, si $P = M_kM_{k-1}\cdots M_1$, entonces PA es la matriz que contiene el resultado de los intercambios de filas realizados a la matriz A para obtener una forma escalonada mediante el método de

Gauss. Luego, si a la matriz PA se le aplica el método para obtener la factorización LU, no se necesitan intercambios de renglones y, entonces,

$$PA = LU.$$

Dado que la matriz P es la matriz identidad con algunas de sus filas permutadas, la matriz P es ortogonal, luego invertible y además $P^{-1} = P^t$ (cfr. teorema 4.14 y definición 4.10, pág. 282).

► **Ejemplo 8.11** Encontrar una matriz de permutación P, una matriz triangular inferior L y una matriz triangular superior U tales que

$$PA = LU$$

si

$$A = \begin{bmatrix} 1 & -2 & 1 & 3 \\ 2 & -4 & 1 & 4 \\ 1 & -2 & 1 & 2 \\ 3 & -5 & 2 & 7 \end{bmatrix}$$ ◄

Solución Llevemos, mediante el método de Gauss, la matriz A a forma escalonada:

$$\begin{bmatrix} 1 & -2 & 1 & 3 \\ 2 & -4 & 1 & 4 \\ 1 & -2 & 1 & 2 \\ 3 & -5 & 2 & 7 \end{bmatrix} \xleftrightarrow[\substack{R_2 \leftrightarrow R_2 - 2R_1 \\ R_3 \leftrightarrow R_3 - R_1 \\ R_4 \leftrightarrow R_4 - 3R_1}]{} \begin{bmatrix} 1 & -2 & 1 & 3 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & -2 \end{bmatrix}$$

$$\xleftrightarrow[R_2 \leftrightarrow R_4]{} \begin{bmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & -2 \end{bmatrix}$$

$$\xleftrightarrow[R_3 \leftrightarrow R_4]{} \begin{bmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Entonces la matriz de permutación se obtiene haciendo los intercambios de renglón $R_2 \leftrightarrow R_4$ y $R_3 \leftrightarrow R_4$ a la matriz identidad; esto es,

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Por tanto, las operaciones de renglón que se realizarían a la matriz PA serían $R_2 \leftrightarrow R_2 - 3R_1$, $R_2 \leftrightarrow R_2 - 2R_1$, $R_4 \leftrightarrow -R_1$. Así,

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

y

$$U = \begin{bmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Comprobación:

$$PA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 & 3 \\ 2 & -4 & 1 & 4 \\ 1 & -2 & 1 & 2 \\ 3 & -5 & 2 & 7 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 & 3 \\ 3 & -5 & 2 & 7 \\ 2 & -4 & 1 & 4 \\ 1 & -2 & 1 & 2 \end{bmatrix}$$

$$LU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 & 3 \\ 3 & -5 & 2 & 7 \\ 2 & -4 & 1 & 4 \\ 1 & -2 & 1 & 2 \end{bmatrix}. \quad ✔$$

Naturalmente no toda matriz tiene factorización LU mediante el método de Gauss; por ejemplo la matriz

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -2 & 2 & -2 \\ 1 & -1 & 1 \end{bmatrix};$$

¿por qué?

Para resolver el sistema $A\vec{x} = \vec{b}$ con factorización LU y matriz de permutación P, se resuelve el sistema $(LU)\vec{x} = \vec{b}$ como hicimos arriba y, si su solución es $\vec{s}$, entonces la del sistema $A\vec{x} = \vec{b}$ es $\vec{x} = P^t\vec{s}$.

Tanto MATLAB como la calculadora HP 50g tienen funciones específicas para calcular la factorización LU de una matriz; consulte los apartados 7.1.8 y 7.2.10. Queda de ejercicio al lector escribir un algoritmo para resolver sistemas lineales $[A\ B]$, $B = [\vec{b}_1\ \vec{b}_2\ \cdots\ \vec{b}_m]$ utilizando factorización LU con matriz de permutación; e implementar este algoritmo en algún lenguaje de programación o en MATLAB para su ejecución en computadora.

8.2.4 Estrategias para pivotar

Los errores de redondeo, que se producen al aplicar el método para resolver sistemas lineales que vimos en el algoritmo 8.1 del apartado anterior, pueden ser muy grandes y, por tanto, dar resultados incorrectos. En el siguiente ejemplo se ilustra esta situación.

► **Ejemplo 8.12** Es fácil ver que el sistema

$$\begin{aligned} 0.0030x_1 &+ 58.13x_2 &= 58.16 \\ 5.292x_1 &- 5.14x_2 &= 47.78 \end{aligned} \tag{8.9}$$

tiene solución exacta $x_1 = 10$ y $x_2 = 1$. Sin embargo, si resolvemos (8.9) por el método de solución de sistemas lineales por eliminación gaussiana con sustitución regresiva (algoritmo 8.1), mediante aritméti-

ca de punto flotante con redondeo a cuatro cifras significativas, tenemos:

$$m_{21} = \frac{5.292}{0.0030} = 1764$$

Entonces, redondeando a 4 cifras significativas,

$$\begin{bmatrix} 0.0030 & 58.13 & 58.16 \\ 5.292 & -5.14 & 47.78 \end{bmatrix} \sim \begin{bmatrix} 0.0030 & 58.13 & 58.16 \\ 0 & 102500 & 102600 \end{bmatrix}.$$

Al hacer sustitución regresiva se obtiene (redondeando a cuatro cifras significativas)

$$\begin{aligned} x_2 &= \frac{102600}{102500} \\ &= 1.001 \end{aligned}$$

y

$$\begin{aligned} x_1 &= \frac{58.16 - 58.13 \cdot 1.001}{0.003} \\ &= \frac{58.16 - 58.19}{0.003} \\ &= -10.00 \end{aligned}$$

El error absoluto para x_2 es

$$|1 - 1.001| = 0.001\,;$$

mientras que para x_1

$$|10 - (-10)| = 20\,.$$

Estrategia de pivote parcial

El problema que presenta el ejemplo anterior radica en que la magnitud, $a_{11} = 0.003$, del pivote[3] inicial en el método es pequeña en comparación con las demás cantidades de la matriz de coeficientes del sistema (8.9). Para evitar —en cierta medida— estas dificultades, es necesario ser más selectivos con el pivote que se elige en cada paso del método de Gauss. Una técnica para este fin, llamada estrategia de pivote parcial, consiste en elegir como pivote, en cada paso del método de Gauss, el elemento de la columna con mayor valor absoluto. En el siguiente ejemplo, aplicamos la estrategia de pivote parcial para resolver el sistema (8.9) del ejemplo 8.12 con aritmética de punto flotante y redondeo a cuatro cifras significativas.

► **Ejemplo 8.13** Dado que el elemento de mayor magnitud en la primera columna de la matriz de coeficientes del sistema (8.9) es $a_{21} = 5.292$, es éste el pivote que se debe elegir para empezar el método de Gauss:

[3]Recordemos que llamamos pivote a un elemento que se elige en una matriz, para que con él, por medio de la operación de renglón $R_i \leftrightarrow \alpha R_i + \beta R_j$, se pueda transformar en ceros los elementos que están por debajo y/o encima de él para obtener una matriz equivalente (cfr. 1.2.4).

$$\begin{bmatrix} 0.0030 & 58.13 & 58.16 \\ 5.292 & -5.14 & 47.78 \end{bmatrix} \sim \begin{bmatrix} \boxed{5.292} & -5.14 & 47.78 \\ 0.0030 & 58.13 & 58.16 \end{bmatrix}.$$

Entonces, al redondear a cuatro cifras significativas,

$$\begin{aligned} m_{21} &= \frac{0.0030}{5.292} \\ &\approx 0.5669 \times 10^{-3} \end{aligned}$$

Luego

$$\begin{bmatrix} \boxed{5.292} & -5.14 & 47.78 \\ 0.0030 & 58.13 & 58.16 \end{bmatrix} \sim \begin{bmatrix} 5.292 & -5.14 & 47.78 \\ 0.0000 & 58.13 & 58.13 \end{bmatrix}$$

implica, al hacer sustitución regresiva,

$$x_2 = 1$$

y

$$\begin{aligned} x_1 &= \frac{47.78 + 5.14}{5.292} \\ &= 10. \end{aligned}$$

los cuales son los valores exactos de la solución del sistema (8.9).◄

Las modificaciones al algoritmo 8.1 que se deben realizar para implementar la estrategia de pivote parcial son mínimas y están contenidas en el siguiente método.

Método de solución de sistemas lineales por eliminación gaussiana con sustitución regresiva y estrategia de pivote parcial

Algoritmo 8.3

Entrada: *La matriz aumentada $A = [A_1\ \vec{b}]$ del sistema (8.3), donde $A_1 = [a_{ij}]$ es la matriz de coeficientes y $\vec{b} = [a_{1,n+1}\ a_{2,n+1}\ \cdots\ a_{n,\,n+1}]^t$ es el término independiente.*
Salida: *La solución $x_1\ x_2\ \cdots\ x_n$ del sistema (8.3) o la leyenda* "El sistema no tiene solución única".

Para cada $j = 1, \ldots, n-1$, de manera recurrente, se genera una matriz $A^{(j)} = [a^{(j)}_{\alpha\beta}]$ equivalente por filas a la matriz A, donde la notación $a^{(0)}_{\alpha\beta}$ representa la componente $a_{\alpha\beta}$ de A, de la siguiente manera (pasos 1 a 4):

Paso 1. *Se elige el menor subíndice $k \geq j$ de la columna j tal que*

$$\left|a^{(j-1)}_{kj}\right| = \max_{j \leq i \leq n} \left|a^{(j-1)}_{ij}\right|.$$

Paso 2. *Si $\left|a^{(j-1)}_{kj}\right| = 0$; es decir, si $a^{(j-1)}_{ij} = 0$ para todo $i \geq j$, se interrumpe la generación de la matriz $A^{(j)}$ y se despliega la salida* "El sistema no tiene solución única".

Paso 3. *Si* $\left|a_{kj}^{(j-1)}\right| \neq 0$, *el proceso continúa. Se calcula*

$$m_{ij}^{(j)} = \frac{a_{ij}^{(j-1)}}{a_{kj}^{(j-1)}}$$

$i = j, j+1, \ldots, n,\ i \neq k$

Paso 4. *Mediante la operación de renglón*

$$R_i \leftrightarrow R_i - m_{ij} R_k$$

se transforman en ceros todos los elementos de la matriz $A^{(j-1)}$ *que están en las filas* $i = j, j+1, \ldots, n,\ i \neq k$; *a la matriz resultante se le aplica la operación de renglón* $R_j \leftrightarrow R_k$ *para obtener la matriz* $A^{(j)}$.

Paso 5. *Si el proceso continúa sin interrupción hasta* $j = n$; *es decir,* $a_{nn}^{(n-1)} \neq 0$, *sean*

$$x_n = \frac{a_{n,\,n+1}^{(n-1)}}{a_{nn}^{(n-1)}}$$

y

$$x_i = \frac{1}{a_{ii}^{(n-1)}}\left[a_{i,\,n+1}^{(n-1)} - \sum_{j=i+1}^{n} a_{ij}^{(n-1)} x_j\right]$$

para $i = n-1, \ldots, 1$. *Entonces la salida es la matriz fila*

$$[x_1\ x_2\ \cdots\ x_n]$$

Programa en MATLAB para el método de Gauss con estrategia de pivote parcial

La figura 8-6 contiene el guión en MATLAB de la función `pivote_parcial`. Este programa esencialmente es el programa `sistemas_gauss` (cfr. figura 8-3) con las modificaciones pertinentes para realizar la estrategia de pivote parcial. Es conveniente que el estudiante siga paso a paso las instrucciones, con alguna entrada concreta, para que pueda entender las partes nuevas que conforman esta función. El programa es muy elemental y tiene como objetivo estimular al lector para que intente mejorarlo.

Estrategia de pivote parcial escalado

Aunque la estrategia de pivote parcial es en general eficiente, para ciertos sistemas puede resultar inadecuada. El siguiente ejemplo contiene uno de estos casos.

► **Ejemplo 8.14** Resolver el sistema lineal

$$\begin{array}{rcrcr} 30x_1 & + & 581300x_2 & = & 581600 \\ 5.292x_1 & - & 5.14x_2 & = & 47.78 \end{array} \tag{8.10}$$

con aritmética de punto flotante y redondeo a cuatro cifras significativas, empleando la estrategia de pivote parcial.◄

```
function M=pivote_parcial(A)
%***pivote_parcial(A)***
%Este programa resuelve, por el método de
%Gauss y utilizando la estrategia de pivote
%parcial, el sistema cuadrado
%Bx=b, donde A=[B b] es la matriz ampliada
%del sistema. El programa emite un mensaje
%cuando el sistema tiene una infinidad
%de soluciones o es inconsistente y da
%como resultado la solución:
%    x1 x2 ... xn
%cuando es única.
%Utiliza las funciones primera_entrada_nonula,
%sust_regresiva, intercambiofilas y sumafilas
n=size(A,1);
A1=A;
mu=1:n;
for j=1:n
   X=A1(:,j)';
   X1=X.*(mu>=j);
   M=max(abs(X1));
   Y=X1.*(X1==-M | X1==M);
   k=primera_entrada_nonula(Y,j);
   if k~=0
      A1=intercambiofilas(A1,k,j);
   for i=j+1:n
   m=-A1(i,j)/X(k);
   A1=sumafilas(A1,i,j,1,m);
   end
   M1=A1;
else
   M1=0;
   break
end
end
if M1==0
   M='El sistema no tiene solución única';
else
   M=sust_regresiva(M1);
end
```

Figura 8-6 • Función pivote_parcial.

Solución El sistema (8.10) es equivalente al sistema (8.9) del ejemplo 8.12, pues la primera ecuación de (8.10) es la primera ecuación de (8.9) multiplicada por el escalar 10^4; así que también tiene la solución exacta $x_1 = 10$ y $x_2 = 1$. Al utilizar la estrategia de pivote parcial se obtiene:

$$\begin{bmatrix} 30 & 581300 & 581600 \\ 5.292 & -5.14 & 47.78 \end{bmatrix} \sim \begin{bmatrix} 30 & 581300 & 581600 \\ 0 & -102500 & -102600 \end{bmatrix}$$

y, por tanto,

$$\begin{aligned} x_1 &= 1.001 \\ x_2 &= -10 \end{aligned}$$

Que resulta la misma aproximación extremadamente burda del ejemplo 8.12. ✔

Una forma para resolver sistemas con dificultades como las que presenta el ejemplo anterior, es la estrategia de **pivote parcial escalado**, que enunciamos a continuación.

Método de solución de sistemas lineales por eliminación gaussiana con sustitución regresiva y estrategia de pivote parcial escalado

Algoritmo 8.4

Entrada: *La matriz aumentada* $A = [A_1\ \vec{b}]$ *del sistema (8.3), donde* $A_1 = [a_{ij}]$ *es la matriz de coeficientes y* $\vec{b} = [a_{1,n+1}\ a_{2,n+1}\ \cdots\ a_{n,n+1}]^t$ *es el término independiente.*

Salida: *La solución* $x_1\ x_2\ \cdots\ x_n$ *del sistema (8.3) o la leyenda* "El sistema no tiene solución única".

Paso 0. *Sean* F_i, $i = 1, 2, \ldots, n$, *las filas de la matriz de coeficientes* A_1. *Para cada* $i = 1, 2, \ldots, n$ *se calcula*

$$s_i = \|F_i\|_\infty = \max_{1 \le j \le n} |a_{ij}|.$$

Si $s_i = 0$ *para algún i, entonces se despliega la información* "El sistema no tiene solución única".

Supongamos que $s_i \neq 0 \; \forall \; i$. *Para cada* $j = 1, \ldots, n-1$, *se genera de manera recurrente una matriz* $A^{(j)} = [a_{\alpha\beta}^{(j)}]$ *—donde la notación* $a_{\alpha\beta}^{(0)}$ *representa la componente* $a_{\alpha\beta}$ *de* A*— equivalente por filas a la matriz* A*, de la siguiente manera (pasos 1 a 4):*

Paso 1. *Se elige el menor subíndice* $k \geq j$ *de la columna* j *tal que*

$$\frac{|a_{kj}^{(j-1)}|}{s_k} = \max_{j \leq i \leq n} \frac{|a_{ij}^{(j-1)}|}{s_i}$$

Paso 2. *Si* $|a_{kj}^{(j-1)}| = 0$*; es decir, si* $a_{ij}^{(j-1)} = 0$ *para todo* $i \geq j$*, se interrumpe la generación de la matriz* $A^{(j)}$ *y se despliega la salida* "El sistema no tiene solución única".

Paso 3. *Si* $|a_{kj}^{(j-1)}| \neq 0$*, el proceso continúa. Se calcula*

$$m_{ij}^{(j)} = \frac{a_{ij}^{(j-1)}}{a_{kj}^{(j-1)}}$$

$i = j, j+1, \ldots, n$, $i \neq k$

Paso 4. *Mediante la operación de renglón*

$$R_i \leftrightarrow R_i - m_{ij} R_k$$

se transforman en ceros todos los elementos de la matriz $A^{(j-1)}$ *que están en las filas* $i = j, j+1, \ldots, n$, $i \neq k$*; a la matriz resultante se le aplica la operación de renglón* $R_j \leftrightarrow R_k$ *para obtener la matriz* $A^{(j)}$*; y se realiza el intercambio* $s_j \leftrightarrow s_k$ *al vector* $\vec{s}^t = [s_i]^t$.

Paso 5. *Si el proceso continúa sin interrupción hasta* $j = n$*; es decir,* $a_{nn}^{(n-1)} \neq 0$*, sean*

$$x_n = \frac{a_{n,n+1}^{(n-1)}}{a_{nn}^{(n-1)}}$$

y

$$x_i = \frac{1}{a_{ii}^{(n-1)}} \left[a_{i,n+1}^{(n-1)} - \sum_{j=i+1}^{n} a_{ij}^{(n-1)} x_j \right]$$

para $i = n-1, \ldots, 1$. *Entonces la salida es la matriz fila*

$$[x_1 \; x_2 \; \cdots \; x_n]$$

▶ **Ejemplo 8.15** Resolver el ejemplo 8.14 con la estrategia de pivote parcial escalado.◀

Solución Para este caso la matriz ampliada del sistema (8.10) es

$$\begin{bmatrix} 30 & 581300 & 581600 \\ 5.292 & -5.14 & 47.78 \end{bmatrix}.$$

Entonces

$$\begin{aligned} s_1 &= \|(30, 58130)\|_\infty \\ &= \max\{|30|, |58130|\} \\ &= 58130 \end{aligned}$$

y

$$\begin{aligned} s_2 &= \|(5.292, -5.14)\|_\infty \\ &= \text{máx}\{|5.292|, |-5.14|\} \\ &= 5.292. \end{aligned}$$

$$\begin{aligned} \text{máx}\left\{\frac{|a_{11}|}{s_1}, \frac{|a_{21}|}{s_2}\right\} &= \text{máx}\left\{\frac{30}{581300}, \frac{5.292}{5.292}\right\} \\ &= 1, \end{aligned}$$

luego $k = 2$ y, por tanto, al aplicar los pasos 1 a 4 del algoritmo 8.4, se obtiene

$$\begin{aligned} \begin{bmatrix} 30 & 581300 & 581600 \\ \boxed{5.292} & -5.14 & 47.78 \end{bmatrix} &\sim \begin{bmatrix} 0 & 581300 & 581300 \\ 5.292 & -5.14 & 47.78 \end{bmatrix} \\ &\sim \begin{bmatrix} 5.292 & -5.14 & 47.78 \\ 0 & 581300 & 581300 \end{bmatrix}. \end{aligned}$$

Al realizar sustitución regresiva,

$$x_2 = 1$$

y

$$\begin{aligned} x_1 &= \frac{47.78 + 5.14}{5.292} \\ &= \frac{52.92}{5.292} \\ &= 10 \end{aligned}$$

que es la solución exacta del sistema (8.10). ✔

▶ **Ejemplo 8.16** Resolver el sistema

$$\begin{aligned} 3.11x_1 &- 5.22x_2 + 0.822x_3 = 3.02 \\ 4.11x_1 &+ 11.3x_2 - 1.14x_3 = -4.10 \\ 1.08x_1 &+ 0.897x_2 + 0.834x_3 = 3.45 \end{aligned}$$

con aritmética de punto flotante redondeando a 3 cifras significativas, mediante la estrategia de pivote parcial escalado del algoritmo 8.4.◀

Solución La matriz ampliada que corresponde al sistema es

$$A = \begin{bmatrix} 3.11 & -5.22 & 0.822 \\ 4.11 & 11.3 & -1.14 \\ 1.08 & 0.897 & 0.834 \end{bmatrix}.$$

Entonces, en este caso, si F_i, $i = 1,2,3$ son las filas de la matriz A_1 —la matriz de coeficientes del sistema—,

$$s_1 = \|F_1\|_\infty = 5.22,$$
$$s_2 = \|F_2\|_\infty = 11.3,$$
$$s_3 = \|F_3\|_\infty = 1.08.$$

Para $j = 1$:

$$\text{máx}\left\{\frac{|a_{11}|}{s_1}, \frac{|a_{21}|}{s_2}, \frac{|a_{31}|}{s_3}\right\} = \text{máx}\left\{\frac{3.11}{5.22}, \frac{4.11}{11.3}, \frac{1.08}{1.08}\right\}$$
$$= 1$$

por lo que $k = 3$. Entonces, ya que ($m_{11} = 2.88$, $m_{21} = 3.81$)

$$\begin{bmatrix} 3.11 & -5.22 & 0.822 & 3.02 \\ 4.11 & 11.3 & -1.14 & -4.10 \\ \boxed{1.08} & 0.897 & 0.834 & 3.45 \end{bmatrix} \sim \begin{bmatrix} 0 & -7.80 & -1.58 & -6.92 \\ 0 & 7.88 & -4.32 & -17.2 \\ 1.08 & 0.897 & 0.834 & 3.45 \end{bmatrix},$$

se tiene

$$A^{(1)} = \begin{bmatrix} 1.08 & 0.897 & 0.834 & 3.45 \\ 0 & 7.88 & -4.32 & -17.2 \\ 0 & -7.80 & -1.58 & -6.92 \end{bmatrix}.$$

El intercambio $s_1 \leftrightarrow s_3$ produce $\vec{s} = [1.08\ \ 11.3\ \ 5.22]$.

Para $j = 2$:

$$\text{máx}\left\{\frac{|a_{22}^{(1)}|}{s_2}, \frac{|a_{32}^{(1)}|}{s_3}\right\} = \text{máx}\left\{\frac{7.88}{11.3}, \frac{7.80}{5.22}\right\}$$
$$= \frac{7.80}{5.22}$$

por lo que $k = 3$. Entonces, ya que ($m_{22} = -1.01$)

$$\begin{bmatrix} 1.08 & 0.897 & 0.834 & 3.45 \\ 0 & 7.88 & -4.32 & -17.2 \\ 0 & \boxed{-7.80} & -1.58 & -6.92 \end{bmatrix} \sim \begin{bmatrix} 1.08 & 0.897 & 0.834 & 3.45 \\ 0 & 0 & -5.92 & -24.2 \\ 0 & -7.80 & -1.58 & -6.92 \end{bmatrix},$$

se tiene

$$A^{(2)} = \begin{bmatrix} 1.08 & 0.897 & 0.834 & 3.45 \\ 0 & -7.80 & -1.58 & -6.92 \\ 0 & 0 & -5.92 & -24.2 \end{bmatrix}.$$

Sustitución regresiva:

$$x_3 = \frac{24.2}{5.92}$$
$$= 4.09$$

$$x_2 = \frac{-6.92 + 1.58 \cdot 4.09}{-7.80}$$
$$= 0.059$$

$$x_1 = \frac{3.45 - 0.897 \cdot 0.059 - 0.834 \cdot 4.09}{1.08}$$
$$= -0.0119 \quad ✔$$

Programa en MATLAB para el método de Gauss con estrategia de pivote parcial escalado

La figura 8-7 contiene el guión del programa en MATLAB de la función pivote_parcial_escalado. Las instrucciones para su uso vienen dadas en los comentarios iniciales. Es conveniente que el lector pruebe, con datos concretos, los comandos de este programa para que por sí mismo comprenda su funcionamiento.

```
function z=pivote_parcial_escalado(A)
%***pivote_parcial_escalado(A)***
%La función pivote_parcial_escalado resuelve
%el sistema lineal A=[A1 b], por medio
%de pivote parcial escalado;
%su salida es la solución del sistema
%o el mensaje de que no tiene solución única.
%Utiliza las funciones intercambiofilas, sumafilas
%y sust_regresiva.
n=size(A,1);
for i=1:n
   s(i)=max(abs(A(i,1:n)));
end
if sum(s~=zeros(1,n))<n
   p=0;
else
   for j=1:n
      [max k]=max(abs(A(j:n,j)'./s(1,j:n)));
      if max==0
         p=0;
         break
      else
         k=k+j-1;
         A=intercambiofilas(A,k,j);
         s=intercambiofilas(s',k,j)';
         for i=j+1:n
             m=-A(i,j)/A(j,j);
             A=sumafilas(A,i,j,1,m);
         end
      end
      p=A;
   end
end

if p==0
     z='el sistema no tiene solución única';
  else
z=sust_regresiva(p);
 end
```

Figura 8-7 • Función pivote_parcial_escalado.

❍ **Nota 8.2** Los programas que hasta aquí se han escrito y los que se desarrollarán más adelante en este capítulo, tienen como objetivo dar ideas generales de la programación, para ejecutarse en una computadora, de los algoritmos que se emplean en cada segmento. Se ha utilizado el paquete MATLAB porque ya se estudió en el capítulo anterior; sin embargo, se pretende que los guiones puedan servir como base al lector para escribir, si así lo desea, estos programas en algún otro lenguaje de programación como C o C++, etc. Las razones principales de hacerlo así, en lugar de redactar pseudocódigos, son el proveer, al menos en un software, programas que el lector —con pocas bases en programación— pueda utilizar como apoyo computacional con un mínimo de conocimientos en este campo y que el presente texto en este aspecto también sea autocontenido. Por otra parte, los programas en MATLAB de este capítulo, son todos muy elementales; de hecho se han pensado más para exponer facetas al programar en este paquete, que en aspectos de optimización en tiempos de ejecución, elegancia, etc. Por tanto, todos esos programas son susceptibles a mejoras, por lo que se invita al lector a perfeccionarlos.

Estrategia de pivote total

Existe una estrategia para pivotear que suele ser más eficiente que las estrategias precedentes para minimizar errores de redondeo. Se llama estrategia **de pivote total** (o **completo**); consiste esencialmente en elegir como pivote, en cada paso del método de Gauss, al elemento con mayor valor absoluto de una submatriz de la matriz original. Por tanto, esta técnica puede involucrar, además del intercambio de renglones, intercambio de columnas. Ilustramos, en el siguiente ejemplo, la estrategia de pivote total aplicada a un sistema lineal específico.

► **Ejemplo 8.17** Resolver, con aritmética de punto flotante redondeando a 4 cifras significativas, el sistema lineal

$$\begin{array}{rcrcrcr} 21.23x_1 & + & 18.23x_2 & - & 40.28x_3 & = & -63.15 \\ -2.21x_1 & + & 33.00x_2 & - & 1.12x_3 & = & 60.43 \\ -1.15x_1 & - & 10.01x_2 & - & 2.29x_3 & = & -28.04 \end{array} \tag{8.11}$$

Solución La matriz ampliada del sistema (8.11) es

$$A = [A_1\ \vec{b}] = \begin{bmatrix} 21.23 & 18.23 & -40.28 & -63.15 \\ -2.21 & 33.00 & -1.12 & 60.43 \\ -1.15 & -10.01 & -2.29 & -28.04 \end{bmatrix}$$

1. Calculamos

$$\|A_1\|_\infty = \max_{\substack{1\leq i\leq 3\\ 1\leq j\leq 3}} |a_{ij}| = 40.28$$

2. Intercambiamos las columnas 3 y 1; es decir, $K_1 \leftrightarrow K_3$, porque $\|A_1\|_\infty$ se alcanza en la componente a_{13} de la matriz A_1:

$$\begin{array}{c} \begin{array}{cccc} x_3 & x_2 & x_1 & \end{array} \\ \begin{bmatrix} -40.28 & 18.23 & 21.23 & -63.15 \\ -1.12 & 33.00 & -2.21 & 60.43 \\ -2.29 & -10.01 & -1.15 & -28.04 \end{bmatrix} \end{array}$$

3. Aplicamos el método de Gauss a la matriz precedente tomando como primer pivote a -40.28 ($m_{21} = .02781$, $m_{31} = 0.05685$):

$$\begin{array}{c} \begin{array}{cccc} x_3 & x_2 & x_1 & \end{array} \\ \begin{bmatrix} \boxed{-40.28} & 18.23 & 21.23 & -63.15 \\ -1.12 & 33.00 & -2.21 & 60.43 \\ -2.29 & -10.01 & -1.15 & -28.04 \end{bmatrix} \end{array} \sim \begin{array}{c} \begin{array}{cccc} x_3 & x_2 & x_1 & \end{array} \\ \begin{bmatrix} -40.28 & 18.23 & 21.23 & -63.15 \\ 0 & 32.49 & -2.8 & 62.19 \\ 0 & -11.05 & -2.357 & -24.45 \end{bmatrix} \end{array}$$

$$= A^{(1)}$$

4. Se calcula

$$\max_{\substack{2\leq i\leq 3\\ 2\leq j\leq 3}} |a_{ij}^{(1)}| = \left\| \begin{bmatrix} 32.49 & -2.8 \\ -11.05 & -2.357 \end{bmatrix} \right\|_\infty$$

$$= 32.49$$

5. Aplicamos el método de Gauss a la matriz precedente tomando como segundo pivote a 32.49 ($m_{32} = -0.3401$):

$$\begin{array}{cccc} x_3 & x_2 & x_1 & \\ \left[\begin{array}{cccc} -40.28 & 18.23 & 21.23 & -63.15 \\ 0 & \boxed{32.49} & -2.8 & 62.19 \\ 0 & -11.05 & -2.357 & -24.45 \end{array}\right] \end{array} \sim \begin{array}{c} \begin{array}{ccc} x_3 & x_2 & x_1 \end{array} \\ \left[\begin{array}{cccc} -40.28 & 18.23 & 21.23 & -63.15 \\ 0 & 32.49 & -2.8 & 62.19 \\ 0 & 0 & -3.309 & -3.3 \end{array}\right] \end{array}$$

6. Se hace sustitución regresiva:

$$\begin{aligned} x_1 &= 0.9973 \\ x_2 &= \frac{62.19 + 2.8 \cdot 0.9973}{32.49} \\ &= 2 \\ x_3 &= \frac{-63.15 - 21.23 \cdot 0.9973 - 18.23 \cdot 2}{-40.28} \\ &= 2.999 \quad ✔ \end{aligned}$$

La solución exacta del sistema lineal del ejemplo anterior es $x_1 = 1$, $x_2 = 2$ y $x_3 = 3$, como fácilmente se puede comprobar. Queda de ejercicio al lector redactar el correspondiente algoritmo para la técnica de pivote total y programar una función en MATLAB o en algún otro lenguaje de programación que aplique esta estrategia para resolver sistemas lineales cuadrados.

8.3 Métodos iterativos

En esta sección estudiaremos tres métodos iterativos para aproximar soluciones de sistemas lineales cuadrados: el de Jacobi, el de Richardson y el de Gauss-Seidel. En este tipo de procedimientos se comienza con un valor inicial de aproximación $\vec{x}_0$ y, de manera iterativa, se calculan aproximaciones sucesivas que, bajo ciertas condiciones, convergen a la solución exacta del sistema lineal. Las técnicas iterativas rara vez se aplican a sistemas lineales de orden bajo, pues el tiempo de ejecución para alcanzar un grado de exactitud satisfactoria suele ser mucho mayor que el tiempo que se requiere cuando se aplica alguno de los métodos directos que vimos en el apartado precedente. Sin embargo, si un sistema de tamaño relativamente grande, $A\vec{x} = \vec{b}$, tiene un alto porcentaje de entradas $a_{ij} = 0$ en la matriz de coeficientes, entonces los métodos iterativos suelen ser muy eficientes, tanto en tiempo de ejecución como en espacio de memoria en la computadora. Estos sistemas, llamados **sistemas ralos**, se presentan con bastante frecuencia en aplicaciones como son la resolución numérica de circuitos eléctricos, ecuaciones diferenciales parciales, problemas con valores en la frontera y equilibrio estático de estructuras, entre otros. La base de los métodos iterativos de Jacobi, Richardson y Gauss-Seidel es el célebre y fundamental teorema del punto fijo del análisis matemático.

8.3.1 *La teoría de punto fijo y normas matriciales naturales*

Teoría de punto fijo en espacios vectoriales normados completos

El caso particular de esta teoría en el espacio $\mathbb{R}$ tiene una motivación geométrica bastante simple e interesante. Consideremos una función de variable real $y = g(x)$ continua en un intervalo $[a, b]$ y supon-

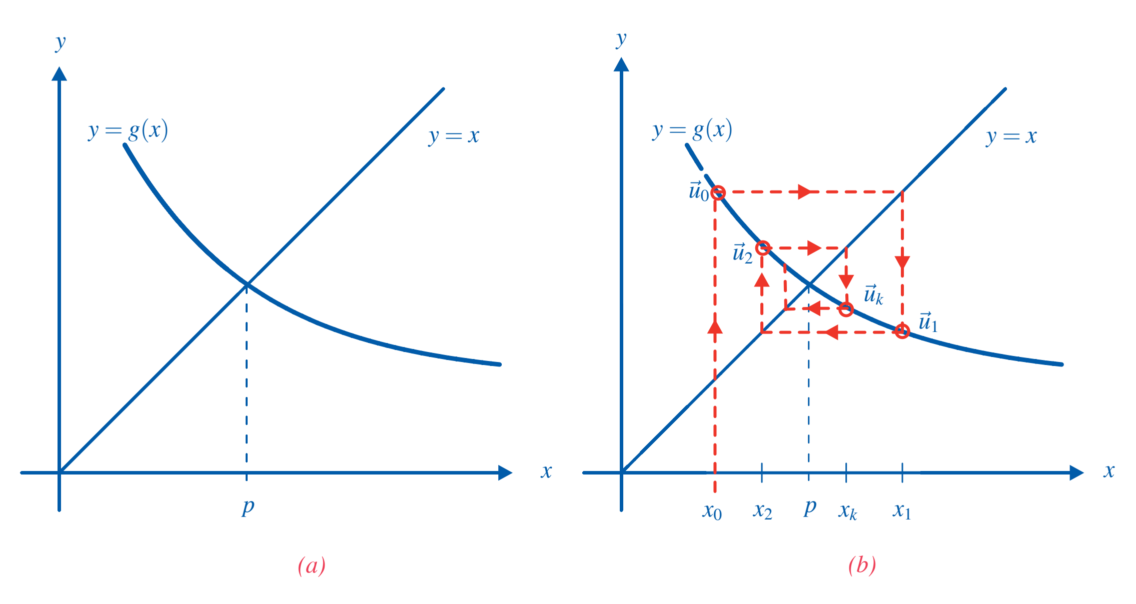


Figura 8-8 •

gamos que se requiere aproximar la solución p (que hemos supuesto existe y es única en el intervalo $[a,b]$) de la ecuación

$$g(x) = x \tag{8.12}$$

Se dice entonces que p es un *punto fijo* de la función g porque $g(p) = p$. La figura 8-8*(a)* ilustra una hipotética función con un único punto fijo p en un intervalo $[a,b]$. Es claro que la solución de (8.12) es la abscisa de la intersección de las gráficas de las funciones $y = g(x)$ y $y = x$, como se muestra en esa figura. Por otra parte, si x_0 es una aproximación inicial de la solución exacta p, entonces, al seguir la trayectoria punteada mostrada en la figura 8-8*(b)*, las abscisas de los puntos $\vec{u}_k = (x_k, g(x_k))$ sobre la gráfica de la función g, tienden a la solución p. Ahora bien, es fácil probar que la abscisa del punto $\vec{u}_1$ es $g(x_0)$; la abscisa del punto $\vec{u}_2$ es $x_2 = g(x_1)$ y, en forma inductiva, la abscisa del punto $\vec{u}_k$ está dada por $x_k = g(x_{k-1})$. De esta manera se obtiene la sucesión (x_k), definida de acuerdo con la relación de recurrencia

$$x_k = \begin{cases} x_0, & \text{si } k = 0; \\ g(x_{k-1}), & \text{si } k = 1, 2, \ldots \end{cases}$$

que converge a la solución exacta p de la ecuación (8.12); es decir,

$$\lim_{k\to\infty} x_k = p.$$

Obviamente la construcción de la sucesión y su convergencia con la solución de la ecuación son meramente intuitivas y requieren de una formulación rigurosa para su demostración, la cual está contenida en el teorema 8.2.

Definición 8.2 *Sean* **E** *un espacio vectorial normado y* $g : \mathbf{E} \to \mathbf{E}$ *una función. Un* ***punto fijo*** *de g en* **E** *es cualquier vector* $\vec{p} \in \mathbf{E}$ *que satisface*

$$g(\vec{p}) = \vec{p}.$$

► **Ejemplo 8.18** Encontrar los puntos fijos de la función $g:\mathbb{R}\to\mathbb{R}$ definida por

$$g(x)=x^2$$

para todo $x\in\mathbb{R}$.◄

Solución Necesitamos hallar todos los números reales p tales que

$$p^2=p.$$

Lo anterior equivale a calcular las soluciones de la ecuación

$$p(p-1)=0.$$

Luego

$$\begin{aligned} p_1 &= 1,\\ p_2 &= 0 \end{aligned}$$

son todos los puntos fijos de g en $\mathbb{R}$. ✔

Teorema 8.2 **(Teorema del punto fijo)** *Sean* **E** *un espacio vectorial normado completo*[4] *con norma* $\|\cdot\|$ *y* $g:\mathbf{E}\to\mathbf{E}$ *una función. Si existe una constante* $\alpha\in\mathbb{R}$, *con* $0<\alpha<1$, *tal que*

$$\|g(\vec{u})-g(\vec{v})\|\leq\alpha\|\vec{u}-\vec{v}\| \tag{8.13}$$

para todo par de vectores $\vec{u},\vec{v}\in\mathbf{E}$, *se cumple lo siguiente:*

1. *Sea* $\vec{x}_0$ *cualquier vector en* **E**, *entonces la sucesión de punto fijo,* $(\vec{x}_k)$, *definida de acuerdo con la relación de recurrencia*

$$\vec{x}_k=\begin{cases}\vec{x}_0, & \text{si } k=0\\ g(x_{k-1}), & \text{si } k=1,2,\ldots\end{cases} \tag{8.14}$$

converge a un vector $\vec{p}\in\mathbf{E}$; *esto es,*

$$\lim_{k\to\infty}\vec{x}_k=\vec{p}$$

o, equivalentemente,

$$\lim_{k\to\infty}\|\vec{x}_k-\vec{p}\|=0.$$

2. *Si además* g *es continua*[5] *en el espacio* **E**, *entonces* $\vec{p}$ *es un punto fijo de* g *en* **E**.
3. $\vec{p}$ *es el único punto fijo de* g *en* **E**.

DEMOSTRACIÓN ■ 1. Por la hipótesis (8.13)

$$\begin{aligned}\|\vec{x}_2-\vec{x}_1\| &= \|g(\vec{x}_1)-g(\vec{x}_0)\|\\ &\leq \alpha\|\vec{x}_1-x_0\|;\end{aligned}$$

y, por tanto, utilizando nuevamente (8.13),

[4] Un espacio vectorial normado es **completo** o **de Banach**, si toda sucesión, $(\vec{x}_k)$, que es de Cauchy, esto es, $\lim_{\nu,\mu\to\infty}\|\vec{x}_\nu-\vec{x}_\mu\|=0$, es convergente; es decir, existe $\vec{p}\in\mathbf{E}$ tal que $\lim_{k\to\infty}\vec{x}_k=\vec{p}$; o sea, $\lim_{k\to\infty}\|\vec{x}_k-\vec{p}\|=0$. Cfr. Nota 6.1 de la pág 607.

[5] Que g sea una función continua en **E**, significa que $\lim_{\vec{u}\to\vec{u}_0}g(\vec{u})=g(\vec{u}_0)$ para todo $\vec{u}_0\in\mathbf{E}$; es decir, $\lim_{\vec{u}\to\vec{u}_0}\|g(\vec{u})-g(\vec{u}_0)\|=0$. Equivalentemente, para toda sucesión $(\vec{x}_k)$ de vectores tal que $\lim_{k\to\infty}\|\vec{x}_k-\vec{u}_0\|=0$ se tiene $\lim_{k\to\infty}\|g(\vec{x}_k)-g(\vec{u}_0)\|=0$. Cfr. definición 6.8.

$$\begin{aligned}\|\vec{x}_3-\vec{x}_2\| &= \|g(\vec{x}_2)-g(\vec{x}_1)\|\\ &\le \alpha\,\|\vec{x}_2-\vec{x}_1\|\\ &\le \alpha\cdot\alpha\,\|\vec{x}_1-\vec{x}_0\|\\ &= \alpha^2\,\|\vec{x}_1-\vec{x}_0\|\,.\end{aligned}$$

Supongamos que $m>1$ es un entero y se cumple

$$\|\vec{x}_m-\vec{x}_{m-1}\| \le \alpha^{m-1}\,\|\vec{x}_1-\vec{x}_0\| \tag{8.15}$$

Entonces, aplicando esta hipótesis y (8.13) se tiene

$$\begin{aligned}\|\vec{x}_{m+1}-\vec{x}_m\| &= \|g(\vec{x}_m)-g(\vec{x}_{m-1})\|\\ &\le \alpha\,\|\vec{x}_m-\vec{x}_{m-1}\|\\ &\le \alpha\cdot\alpha^{m-1}\,\|\vec{x}_1-\vec{x}_0\|\\ &\le \alpha^m\,\|\vec{x}_1-\vec{x}_0\|\,.\end{aligned}$$

Luego, por inducción, se concluye que (8.15) es válida para todo entero $m\ge 1$. Sea ahora $k\ge 1$ un entero fijo y ν cualquier entero mayor que k. Entonces, por la desigualdad triangular y (8.15),

$$\begin{aligned}\|\vec{x}_\nu-\vec{x}_k\| &= \|\vec{x}_\nu-\vec{x}_{\nu-1}+\vec{x}_{\nu-1}-\vec{x}_{\nu-2}+\cdots-\vec{x}_{k+1}+\vec{x}_{k+1}-\vec{x}_k\|\\ &\le \|\vec{x}_\nu-\vec{x}_{\nu-1}\|+\|\vec{x}_{\nu-1}-\vec{x}_{\nu-2}\|+\cdots+\|\vec{x}_{k+1}-\vec{x}_k\|\\ &\le \alpha^{\nu-1}\,\|\vec{x}_1-\vec{x}_0\|+\alpha^{\nu-2}\,\|\vec{x}_1-\vec{x}_0\|+\cdots+\alpha^k\,\|\vec{x}_1-\vec{x}_0\|\\ &= \alpha^k\,\|\vec{x}_1-\vec{x}_0\|\left(1+\alpha+\cdots+\alpha^{\nu-k-1}\right).\end{aligned}$$

Ya que la serie geométrica $\sum_{m=0}^{\infty}\alpha^m=\frac{1}{1-\alpha}$, pues $0<\alpha<1$, la precedente desigualdad implica

$$\|\vec{x}_\nu-\vec{x}_k\| \le \frac{\alpha^k}{1-\alpha}\,\|\vec{x}_1-\vec{x}_0\| \tag{8.16}$$

Luego, ya que $0<\alpha<1$,

$$0\le \lim_{\nu,k\to\infty}\|\vec{x}_\nu-\vec{x}_k\| \le \lim_{\nu,k\to\infty}\frac{\alpha^k}{1-\alpha}\,\|\vec{x}_1-\vec{x}_0\|=0$$

y, por tanto,

$$\lim_{\nu,k\to\infty}\|\vec{x}_\nu-\vec{x}_k\|=0.$$

Así, la sucesión $(\vec{x}_k)$ es de Cauchy en el espacio normado completo $\mathbf{E}$ y, por ende, converge a un vector en él; es decir, existe $\vec{p}\in\mathbf{E}$ tal que

$$\lim_{k\to\infty}\vec{x}_k=\vec{p}.$$

2. Debido a la continuidad de la función g en el espacio $\mathbf{E}$ se desprende

$$\vec{p}=\lim_{k\to\infty}\vec{x}_k=\lim_{k\to\infty}g(\vec{x}_{k-1})=g(\vec{p}).$$

Por lo que $\vec{p}$ es un punto fijo de g en el espacio $\mathbf{E}$.

3. Sea $\vec{p}_1$ un punto fijo de g en **E** distinto de $\vec{p}$. Entonces

$$\begin{aligned}\|\vec{p}-\vec{p}_1\| &= \|g(\vec{p})-g(\vec{p}_1)\| \\ &\leq \alpha\,\|\vec{p}-\vec{p}_1\|\end{aligned}$$

y ya que $0<\alpha<1$, lo anterior implica

$$\|\vec{p}-\vec{p}_1\| < \|\vec{p}-\vec{p}_1\|\,.$$

Lo cual es una contradicción; por tanto g tiene únicamente el punto fijo $\vec{p}$ en **E**. ■

Una consecuencia inmediata de la demostración del teorema 8.2, es que existe una manera de acotar el error, $\|\vec{x}_k-\vec{p}\|$, que se comete al aproximar el punto fijo $\vec{p}$ con la k-ésima iteración de la sucesión de punto fijo; nos referimos a la desigualdad (8.16). Sin embargo, dada la importancia de este resultado, lo hacemos patente en el siguiente corolario.

Corolario 8.1 **(Cota para el error en iteraciones de punto fijo)** *Sean* **E**, $\|\cdot\|$, g, α, $\vec{p}$ *y* $\vec{x}_0$ *como en el teorema precedente. Entonces la* k*-ésima iteración de la sucesión de punto fijo (8.14) satisface*

$$\|\vec{x}_k-\vec{p}\| \leq \frac{\alpha^k}{1-\alpha}\,\|\vec{x}_1-\vec{x}_0\| \qquad (8.17)$$

❍ **Nota 8.3** 1. Una función g, definida en un espacio normado **E**, que satisface la desigualdad (8.13) para cierta constante $\alpha>0$, se llama **función de Lipschitz** y a cualquier constante $\alpha>0$ que satisface la desigualdad (8.13) se le dice **constante de Lipschitz** para la función g. Si además, como en el caso del teorema del punto fijo, la constante α es menor a uno, diremos que la función g es una **contracción.**

2. La convergencia de la sucesión de punto fijo se da para cualquier valor inicial $\vec{x}_0\in\mathbf{E}$; sin embargo, si en un problema dado se tiene la idea de una primera estimación que es cercana al valor exacto, entonces se debe emplear ésta como punto inicial con el fin de reducir el número de iteraciones.
3. Se dice que la relación (8.17) es **una cota** para el error en la k-ésima iteración de punto fijo.

► **Ejemplo 8.19** Aproximar la solución de la ecuación

$$\frac{1}{10}\cos x - x = 0 \qquad (8.18)$$

utilizando el menor número de iteraciones de punto fijo para un error inferior a 10^{-3}.◄

Solución La figura 8-9 contiene las gráficas de las funciones $y=x$ y $y=\frac{1}{10}\cos x$; en ella se observa que la solución de la ecuación (8.18) es la abscisa del punto de intersección de ambas gráficas; evidentemente el problema planteado equivale a encontrar el punto fijo de la función $g:\mathbb{R}\to\mathbb{R}$, con

$$g(x)=\frac{1}{10}\cos x$$

para todo $x\in\mathbb{R}$. En este caso el espacio subyacente es $\mathbb{R}$ y la norma en él es el valor absoluto. Antes que nada debemos encontrar una constante de Lipschitz para la función g si es que existe. Sean $x,y\in\mathbb{R}$,

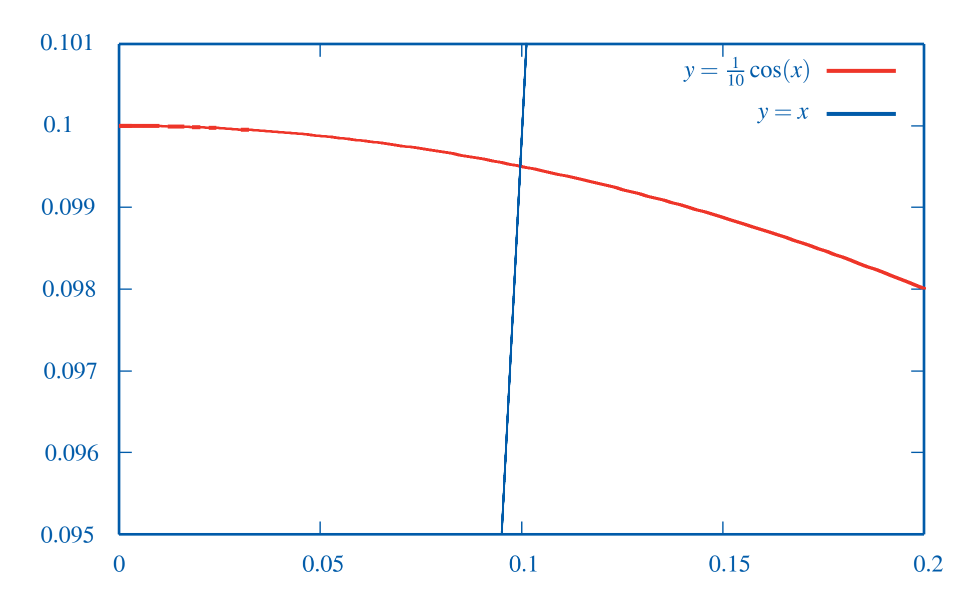


Figura 8-9 •

con $y > x$, un par de números reales cualesquiera. Por el teorema del valor medio (cfr. teorema 6.17) existe $\mu \in [x,y]$, tal que

$$\frac{g(y)-g(x)}{y-x} = g'(\mu);$$

esto es,

$$\frac{\frac{1}{10}\cos(y) - \frac{1}{10}\cos(x)}{y-x} = -\frac{1}{10}\operatorname{sen}\mu$$

y, por tanto,

$$\frac{\left|\frac{1}{10}\cos(y) - \frac{1}{10}\cos(x)\right|}{|y-x|} = \frac{1}{10}|\operatorname{sen}\mu| \leq \frac{1}{10}$$

De donde,

$$|g(x)-g(y)| \leq \frac{1}{10}|x-y| \;\; \forall x,y \in \mathbb{R}.$$

Luego $\alpha = 1/10$ es una constante de Lipschitz para la contracción g. Por el teorema del punto fijo la sucesión

$$x_k = \begin{cases} x_0, & \text{si } k=0 \\ g(x_{k-1}) & \text{si } k=1,2,\ldots \end{cases}$$

converge al único punto fijo de g en $\mathbb{R}$ para cualquier valor inicial x_0; en particular, para $x_0 = 0$. Por (8.17) del corolario 8.1, se tiene entonces

$$
\begin{aligned}
|x_k - p| &\leq \frac{(1/10)^k}{1-1/10}|x_1 - x_0| \\
&= \frac{1}{9\cdot 10^{k-1}}\left|\frac{1}{10}\cos(0) - 0\right| \\
&= \frac{1}{9\cdot 10^k}
\end{aligned}
$$

Puesto que se quiere realizar el menor número de iteraciones para obtener un error inferior a 10^{-3}, basta encontrar el menor entero positivo k tal que

$$\frac{1}{9\cdot 10^k} < 10^{-3}.$$

Esto es, el primer entero positivo k que satisface

$$10^k > \frac{10^3}{9};$$

es decir,

$$k = 3.$$

Finalmente, las tres primeras iteraciones de punto fijo son

$$
\begin{aligned}
x_1 &= g(0) = 0.1 \\
x_2 &= g(x_1) = 0.0995004 \\
x_3 &= g(0.0995004) = 0.0995054
\end{aligned}
$$

Así, la aproximación de la única solución de la ecuación $\frac{1}{10}\cos x = x$, con un error menor a 10^{-3}, es

$$x_3 = 0.0995054. \checkmark$$

La razón de haber elegido $x_0 = 0$ como aproximación inicial en el ejemplo anterior, fue por la comodidad para realizar las operaciones aritméticas que este número representa; sin embargo, en la gráfica de la figura 8-9 se observa que p es un valor más cercano a 0.1 que a 0 y, por tanto, se espera que el número de iteraciones de punto fijo sea menor[6] para alcanzar la precisión deseada si se toma como valor inicial $x_0 = 0.1$. El lector debe notar que en el ejemplo anterior la aproximación y solución de la ecuación $\frac{1}{10}\cos x - x = 0$ (solución que es imposible obtener en forma directa, intente despejar x), difieren en menos de 10^{-3} y esta cota —que se fija previamente para el error— se pudo obtener, gracias al corolario 8.1, sin saber el valor exacto de la solución p. Es decir, no conocemos p, pero podemos asegurar que $|x_3 - p| < 0.001$; en realidad es posible determinar, dado cualquier $\varepsilon > 0$, el menor número de iteraciones k para que $|x_k - p| < \varepsilon$. ¿No es ésta una maravillosa aplicación de las matemáticas puras? El ejemplo anterior es un típico caso donde se alcanzan los objetivos del programa fundamental del *análisis numérico*:

[6] Se puede probar, de manera análoga a como se hizo en 1 del teorema 8.2, bajo las mismas hipótesis, que $\|\vec{x}_k - \vec{p}\| \leq \alpha^k \|\vec{x}_0 - \vec{p}\|$; razón por la que el número de iteraciones disminuye si $\vec{x}_0$ es más cercano a $\vec{p}$.

Dado un problema cuya solución p habita en cierto espacio normado $(\mathbf{E}, \|\cdot\|)$, se requiere encontrar una sucesión (x_k) de elementos del mismo espacio, tal que:

- El valor de x_0 es dado y la sucesión (x_k) está definida en forma recurrente.
- $\text{lím}_{k\to\infty} x_k = p$ respecto a la norma $\|\cdot\|$; esto es, $\text{lím}_{k\to\infty} \|x_k - p\| = 0$.
- Es posible acotar el error; es decir, existe una función $k \mapsto f(k)$ (que tiende a cero cuando k tiende a infinito) tal que

$$\|x_k - p\| \leq f(k) \;\; \forall k.$$

No en todos los casos se puede acotar el error; sin embargo, aun así, es posible determinar —basados en el "sentido común" y apoyados en la convergencia de la sucesión al valor exacto de la solución— criterios para fijar hasta qué valor de k se debe detener el ciclo iterativo con el fin de alcanzar una aproximación suficientemente precisa de p. Por esta razón —y por otras como son las limitaciones computacionales de cada equipo, la naturaleza de cada problema a resolver, errores de redondeo, etc.— es que la experiencia que se adquiere a lo largo de resolver una gran variedad de problemas numéricos es indispensable; a tal grado que para los expertos en métodos numéricos esta disciplina se convierte en algo que ellos mismos consideran un "arte".

Normas matriciales naturales

Aunque en un espacio vectorial de dimensión finita todas las normas son equivalentes y, por tanto, la convergencia es independiente de la norma que se utilice (cfr. teorema 4.22 de 4.2.4), en la teoría de aproximación de soluciones de sistemas lineales conviene trabajar con normas en el espacio $\mathfrak{M}_n(\mathbb{R})$ que satisfagan desigualdades multiplicativas del tipo

$$\|A\vec{x}\|_* \leq \|A\|^* \, \|\vec{x}\|_* \tag{*}$$

donde $\|\cdot\|^*$ es una norma en $\mathfrak{M}_n(\mathbb{R})$ y $\|\cdot\|_*$ es una norma en $\mathbb{R}^n$. En las proposiciones 6.3 y 6.4 del apartado 6.3.4, vimos que

$$\|A\|^* = \max_{1\leq j\leq n} \left\{ \sum_{i=1}^{n} |a_{ij}| \right\}$$

satisface la relación de orden (*) para la norma vectorial $\|\vec{x}\|_* = \|\vec{x}\|_1 = \sum_{i=1}^{n} |x_i|$, si $\vec{x} = (x_1, x_2, \ldots, x_n)$. Existen, sin embargo, muchas otras normas matriciales que provienen de normas vectoriales de manera natural y que cumplen con esa desigualdad multiplicativa. La forma de construir éstas se encuentra plasmada en la siguiente proposición.

Proposición 8.1 *Sean T una matriz cuadrada de orden n y $\|\cdot\|$ una norma vectorial dada en $\mathbb{R}^n$. Entonces*[7]

$$\|T\| = \max_{\substack{\vec{x}\in\mathbb{R}^n \\ \|\vec{x}\|=1}} \|T\vec{x}\| \tag{8.19}$$

es una norma en el espacio de matrices $\mathfrak{M}_n(\mathbb{R})$.

[7] Ya que la aplicación $\vec{x} \mapsto T\vec{x}$ es continua y la esfera de centro $\vec{0}_{\mathbb{R}^n}$ y radio 1 es un conjunto compacto, está garantizado que la función $\vec{x} \mapsto \|T\vec{x}\|$ alcanza un máximo en este conjunto.

DEMOSTRACIÓN 1. Claramente $\|T\| \geq 0$ para toda $T \in \mathfrak{M}_n(\mathbb{R})$.

2. Si $T = \mathcal{O}$,

$$\begin{aligned}\|T\| &= \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} \|T\vec{x}\|\\ &= \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} \left\|\vec{0}_{\mathbb{R}^n}\right\|\\ &= 0.\end{aligned}$$

Supongamos que T es una matriz de orden n y que $\|T\| = 0$, entonces de

$$\max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} \|T\vec{x}\| = 0$$

se desprende que

$$T\vec{x} = 0 \quad \forall \vec{x} \in \mathbb{R}^n,\ \|\vec{x}\| = 1.$$

Sean $\vec{z} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$ y $\vec{x} = \frac{1}{\|\vec{z}\|}\vec{z}$, entonces $\|\vec{x}\| = 1$ y por tanto,

$$\begin{aligned}0 &= T\vec{x}\\ &= \frac{1}{\|z\|}T\vec{z};\end{aligned}$$

de donde

$$T\vec{z} = 0 \quad \forall z \in \mathbb{R}^n.$$

Lo cual implica $T = \mathcal{O}$ (cfr. lema 5.1 del apartado 5.2.1, página 436).

3. Si $\beta \in \mathbb{R}$ y $T \in \mathfrak{M}_n(\mathbb{R})$, entonces

$$\begin{aligned}\|\beta T\| &= \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} \|(\beta T)\vec{x}\|\\ &= \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} |\beta|\,\|T\vec{x}\|\\ &= |\beta| \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} \|T\vec{x}\|\\ &= |\beta|\,\|T\|.\end{aligned}$$

4. Si $T_1, T_2 \in \mathfrak{M}_n(\mathbb{R})$, entonces

$$\begin{aligned}\|T_1 + T_2\| &= \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} \|(T_1 + T_2)\vec{x}\|\\ &= \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} \|T_1\vec{x} + T_2\vec{x}\|\\ &\leq \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} \|T_1\vec{x}\| + \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} \|T_2\vec{x}\|\\ &= \|T_2\| + \|T_2\|.\end{aligned}$$

De 1), 2), 3) y 4) $\|\cdot\|$, definida por (8.19), es una norma en $\mathfrak{M}_n(\mathbb{R})$. ■

Definición 8.3 **(Normas matriciales naturales)** *Sea $\|\cdot\|$ una norma vectorial dada; esto es, una norma en $\mathbb{R}^n$. La norma en el espacio de matrices $\mathfrak{M}_n(\mathbb{R})$, definida por (8.19) en la proposición anterior, se llama la* ***norma matricial natural subordinada*** *a la norma vectorial $\|\cdot\|$. Se acostumbra también decir que la norma natural matricial proviene de la norma vectorial o que es inducida por ella.*

Proposición 8.2 *Sea $\|\cdot\|$ una norma vectorial dada, entonces la norma matricial natural subordinada a ella satisface:*

1.
$$\|T\vec{x}\| \leq \|T\| \, \|\vec{x}\| \tag{8.20}$$
para todo $T \in \mathfrak{M}_n(\mathbb{R})$ y para todo $\vec{x} \in \mathbb{R}^n$.

2.
$$\|AB\| \leq \|A\| \, \|B\| \tag{8.21}$$
para todo par de matrices $A, B \in \mathfrak{M}_n(\mathbb{R})$.

3. *Si I_n es la matriz identidad de orden n, $\|I_n\| = 1$.*

DEMOSTRACIÓN ■ 1. Sea $T \in \mathfrak{M}_n(\mathbb{R})$ y $\vec{x} \in \mathbb{R}^n$. Claramente (8.20) se cumple si $\vec{x} = \vec{0}_{\mathbb{R}^n}$. Supongamos que $\vec{x} \neq \vec{0}_{\mathbb{R}^n}$ y sea $\vec{z} = \frac{1}{\|\vec{x}\|}\vec{x}$; entonces de $\|\vec{z}\| = 1$ y de (8.19) se desprende

$$\begin{aligned} \|T\| &\geq \|T\vec{z}\| \\ &= \left\| T \cdot \frac{1}{\|\vec{x}\|}\vec{x} \right\| \\ &= \frac{1}{\|\vec{x}\|} \|T\vec{x}\| \, ; \end{aligned}$$

de donde

$$\|T\vec{x}\| \leq \|T\| \, \|\vec{x}\| \, .$$

2. Si $A, B \in \mathfrak{M}_n(\mathbb{R})$ y $\vec{x} \in \mathbb{R}^n$ es cualquier vector con $\|\vec{x}\| = 1$, entonces, de (8.20), se tiene

$$\begin{aligned} \|(AB)\vec{x}\| &= \|A(B\vec{x})\| \\ &\leq \|A\| \, \|B\vec{x}\| \\ &\leq \|A\| \, \|B\| \, \|\vec{x}\| \\ &= \|A\| \, \|B\| \, . \end{aligned}$$

y, por tanto,

$$\|AB\| = \max_{\substack{\vec{x} \in \mathbb{R}^n \\ \|\vec{x}\|=1}} \|(AB)\vec{x}\| \leq \|AB\|.$$

3. Ejercicio para el lector:

Un par de normas que se utilizan con mucha frecuencia en métodos numéricos son las normas matriciales naturales que están contenidas en los siguientes dos ejemplos. ■

▶ **Ejemplo 8.20** Sea $\|\cdot\|_\infty$, la norma vectorial cúbica (cfr. ejemplo 4.41); esto es,

$$\|(x_1, x_2, \ldots, x_n)\|_\infty = \max_{1\le i\le n} |x_i|.$$

y sea $\|\cdot\|^{(\infty)}$ su norma matricial subordinada; es decir,

$$\|T\|^{(\infty)} = \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|_\infty=1}} \|T\vec{x}\|_\infty .$$

Entonces, si $T = [t_{ij}]$,

$$\|T\|^{(\infty)} = \max_{1\le i\le n}\left\{\sum_{j=1}^{n}|t_{ij}|\right\} \tag{8.22}$$

esto es, $\|T\|^{(\infty)}$ es el máximo de las sumas de los valores absolutos de las componentes de cada fila de la matriz T.◀

DEMOSTRACIÓN ■ En efecto, si $\|\vec{x}\|_\infty = 1$,

$$\begin{aligned}
\|T\vec{x}\|_\infty &= \max_{1\le i\le n}\left\{\sum_{j=1}^{n}|t_{ij}x_j|\right\}\\
&= \max_{1\le i\le n}\left\{\sum_{j=1}^{n}|t_{ij}||x_j|\right\}\\
&\le \max_{1\le i\le n}\left\{\sum_{j=1}^{n}|t_{ij}|\,\|\vec{x}\|_\infty\right\}\\
&= \|\vec{x}\|_\infty \max_{1\le i\le n}\left\{\sum_{j=1}^{n}|t_{ij}|\right\}\\
&= 1\cdot \max_{1\le i\le n}\left\{\sum_{j=1}^{n}|t_{ij}|\right\}.
\end{aligned}$$

Por tanto,

$$\|T\|^{(\infty)} \le \max_{1\le i\le n}\left\{\sum_{j=1}^{n}|t_{ij}|\right\} \tag{8.23}$$

Por otra parte, si $\vec{u} = (1, 1, \ldots, 1) \in \mathbb{R}^n$, entonces $\|\vec{u}\|_\infty = 1$, por ende,

$$\|T\vec{u}\|_\infty \le \max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|=1}} \|T\vec{x}\|_\infty = \|T\|^{(\infty)} .$$

Luego

$$\max\left(\sum_{j=1}^{n}|t_{1j}|, \sum_{j=1}^{n}|t_{2j}|, \ldots, \sum_{j=1}^{n}|t_{nj}|\right) \le \|T\|^{(\infty)} ;$$

esto es,

$$\max_{1\le i\le n}\left\{\sum_{j=1}|t_{ij}|\right\}\le\|T\|^{(\infty)} \tag{8.24}$$

De (8.23) y (8.24), se desprende (8.22). ■

► **Ejemplo 8.21** Sea $\|\cdot\|_1$, la norma vectorial definida en el ejemplo (4.43); esto es,

$$\|(x_1,x_2,\ldots,x_n)\|_1=\sum_{i=1}^{n}|x_i|.$$

y sea $\|\cdot\|^{(1)}$ su norma matricial subordinada; es decir,

$$\|T\|^{(1)}=\max_{\substack{\vec{x}\in\mathbb{R}^n\\ \|\vec{x}\|_1=1}}\|T\vec{x}\|_1\,.$$

Mostrar que si $T=[t_{ij}]$,

$$\|T\|^{(1)}=\max_{1\le j\le n}\left\{\sum_{i=1}^{n}|t_{ij}|\right\} \tag{8.25}$$

esto es, $\|T\|^{(\infty)}$ es el máximo de las sumas de los valores absolutos de las componentes de cada columna de la matriz T (cfr. proposición 6.3 del apartado 6.3.4).◄

DEMOSTRACIÓN ■ Para simplificar la notación, utilizaremos el símbolo $\mathscr{N}(T)$ con el fin de representar el lado derecho de (8.25); es decir, para cualquier matriz $T=[t_{ij}]\in\mathfrak{M}_n(\mathbb{R})$,

$$\mathscr{N}(T)=\max_{1\le j\le n}\left\{\sum_{i=1}^{n}|t_{ij}|\right\}.$$

Sean $\vec{e}_i=(0,\ldots,0,\overset{i}{1},0,\ldots,0)$, $i=1,2,\ldots,n$, los vectores de la base canónica de $\mathbb{R}^n$ y representemos por $\vec{k}_j$ la columna j de la matriz $T=[t_{ij}]$; o sea, $\vec{k}_j=(t_{1j},t_{2j},\ldots,t_{nj})=[\,t_{1j}\quad t_{2j}\quad\cdots\quad t_{nj}\,]^t$; entonces $\|\vec{e}_i\|_1=1$ y, por tanto, (cfr. ejemplo 1.15, pág. 12)

$$\begin{aligned}\|\vec{k}_j\|_1&=\|T\vec{e}_j\|_1\\&\le\|T\|^{(1)}\end{aligned}$$

para todo $j=1,2,\ldots,n$. Luego

$$\max\left\{\|\vec{k}_1\|_1,\|\vec{k}_1\|_1,\ldots,\|\vec{k}_n\|_1\right\}\le\|T\|^{(1)};$$

es decir,

$$\mathscr{N}(T)\le\|T\|^{(1)} \tag{8.26}$$

Por otra parte, si $\vec{x}=(x_1,x_2,\ldots,x_n)$ es un vector con $\|\vec{x}\|_1=1$, entonces (cfr. ejemplo 1.15, pág. 12)

$$\begin{aligned}\|T\vec{x}\|_1 &= \|x_1\vec{k}_1 + x_2\vec{k}_2 + \cdots + x_n\vec{k}_n\|_1 \\ &\le |x_1|\|\vec{k}_1\| + |x_2|\|\vec{k}_2\| + \cdots + |x_n|\|\vec{k}_n\| \\ &\le |x_1|\mathcal{N}(T) + |x_2|\mathcal{N}(T) + \cdots + |x_n|\mathcal{N}(T) \\ &= (|x_1| + |x_2| + \cdots + |x_n|)\mathcal{N}(T) \\ &= \|\vec{x}\|_1\mathcal{N}(T) \\ &= \mathcal{N}(T);\end{aligned}$$

lo cual implica

$$\|T\|^{(1)} \le \mathcal{N}(T) \tag{8.27}$$

De (8.26) y (8.27) se tiene (8.25). ■

❍ **Nota 8.4** En el capítulo 4, definimos normas que son generalizaciones directas de normas vectoriales en el espacio $\mathbb{R}^n$ al espacio $\mathfrak{M}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$ (cfr. ejemplos 4.13, 4.43 y 4.44) y utilizamos las mismas notaciones para representar ambas normas, por ejemplo, $\|\cdot\|_1$ y $\|\cdot\|_\infty$ denotan tanto las normas vectoriales como sus generalizaciones matriciales. Sin embargo, estas normas matriciales no necesariamente son subordinadas de las normas de las cuales son generalizaciones (cfr. ejercicio 327 del capítulo 4) y, por tanto, no satisfacen forzosamente las desigualdades multiplicativas (8.20) y (8.21). Por esta razón, en este libro hemos utilizado las notaciones $\|\cdot\|^*$ para la norma matricial natural subordinada a la norma vectorial $\|\cdot\|_*$, si es que la última está denotada con algún subíndice y así evitar confusiones en notaciones como $\|\cdot\|_\infty$, que es la generalización de la norma cúbica vectorial, $\|(x_1, x_2, \ldots, x_n)\|_\infty = \text{máx}_{1\le i\le n}|x_i|$, con la norma matricial natural $\|\cdot\|^{(\infty)}$, subordinada[8] a la misma norma vectorial.

Teoría de punto fijo para transformaciones afines

Sea T una matriz cuadrada de orden n y $\vec{c}$ un vector fijo de $\mathbb{R}^n$. Se define $g : \mathbb{R}^n \to \mathbb{R}^n$ como

$$g(\vec{x}) = T\vec{x} + \vec{c}$$

para cada $\vec{x} \in \mathbb{R}^n$. Entonces, se dice que g es una **transformación afín** en el espacio $\mathbb{R}^n$.

En la siguiente proposición se demuestra que toda transformación afín es una función de Lipschitz (cfr. nota 8.3), si la norma matricial subyacente es natural.

Proposición 8.3 *Sean $\|\cdot\|$ una norma en $\mathbb{R}^n$ y $\|\cdot\|$ la norma matricial natural que proviene de ella, T una matriz cuadrada de orden n y $\vec{c}$ un vector del espacio $\mathbb{R}^n$. Entonces la transformación afín*

$$g(\vec{x}) = T\vec{x} + \vec{c}$$

es una función de Lipschitz. Específicamente

$$\|g(\vec{x}) - g(\vec{y})\| \le \|T\|\,\|\vec{x} - \vec{y}\|$$

para todo par de vectores $\vec{x}, \vec{y} \in \mathbb{R}^n$.

[8] Los paréntesis son únicamente para enfatizar que la notación es un supraíndice y no un exponente.

DEMOSTRACIÓN ■ Para todo $\vec{x},\vec{y}\in\mathbb{R}^n$, se tiene

$$\begin{aligned}\|g(\vec{x})-g(\vec{y})\| &= \|T\vec{x}+\vec{c}-T\vec{y}-\vec{c}\|\\ &= \|T\vec{x}-T\vec{y}\|\\ &= \|T(\vec{x}-\vec{y})\|\end{aligned}$$

y ya que la norma matricial es natural, se desprende —de (8.20) de la proposición 8.2—,

$$\|g(\vec{x})-g(\vec{y})\| \le \|T\|\,\|\vec{x}-\vec{y}\|\,. \quad ■$$

Ya que todo espacio vectorial de dimensión finita es completo y cualquier operador lineal en él es una función continua, el teorema de punto fijo y su primer corolario tienen un par de consecuencias inmediatas a la teoría de punto fijo para sistemas afines; las hacemos patentes a continuación.

Teorema 8.3 **(Teorema del punto fijo para transformaciones afines)** *Sean $\|\cdot\|$ una norma en $\mathbb{R}^n$ y $\|\cdot\|$ su norma matricial natural subordinada, T una matriz cuadrada de orden n y $\vec{c}$ un vector del espacio $\mathbb{R}^n$. Consideremos la transformación afín en $\mathbb{R}^n$*

$$g(x)=T\vec{x}+\vec{c}.$$

Si $\|T\|<1$, se cumple lo siguiente:

1. *Dado cualquier vector $\vec{x}_0$ en* **E**, *la sucesión de punto fijo definida de acuerdo con la relación de recurrencia*

$$\vec{x}_k=\begin{cases}\vec{x}_0, & \text{si } k=0\\ g(\vec{x}_{k-1}), & \text{si } k=1,2,\dots\end{cases} \tag{8.28}$$

   converge a un vector $\vec{p}\in$ **E**; *esto es,*

$$\lim_{k\to\infty}\vec{x}_k=\vec{p}$$

   o, equivalentemente,

$$\lim_{k\to\infty}\|\vec{x}_k-\vec{p}\|=0.$$

2. *$\vec{p}$ es un punto fijo de g en* **E**.
3. *$\vec{p}$ es el único punto fijo de g en* **E**.

DEMOSTRACIÓN ■ Ya que la norma matricial es natural, del corolario 8.3, la transformación afín $g(\vec{x})=T\vec{x}+\vec{c}$ es una función de Lipschitz; de hecho,

$$\|g(\vec{x})-g(\vec{y})\| \le \|T\|\,\|\vec{x}-\vec{y}\|\ \forall \vec{x},\vec{y}\in\mathbb{R}^n.$$

y ya que $\|T\|<1$, los incisos 1, 2 y 3 son consecuencias de los incisos 1, 2 y 3 del teorema 8.2; pues la función g es continua al ser la suma de una transformación lineal y una constante en el espacio de dimensión finita —y, por ende completo— $\mathbb{R}^n$. ■

Obviamente se tiene la versión particular con transformaciones afines del corolario (8.1) para acotar el error que se comete en la k-ésima iteración de punto fijo para transformaciones afines.

Corolario 8.2 **(Cota para el error en iteraciones de punto fijo para transformaciones afines)** *Sean $\|\cdot\|$ una norma en $\mathbb{R}^n$ y $\|\cdot\|$ su norma matricial natural subordinada, T una matriz cuadrada de orden n y $\vec{c}$ un vector del espacio $\mathbb{R}^n$. Consideremos la transformación afín en $\mathbb{R}^n$*

$$g(x) = T\vec{x} + \vec{c}.$$

Si $\|T\| < 1$, entonces la k-ésima iteración de la sucesión de punto fijo (8.28) satisface

$$\|\vec{x}_k - \vec{p}\| \leq \frac{\|T\|^k}{1 - \|T\|}\|\vec{x}_1 - \vec{x}_0\| \qquad (8.29)$$

8.3.2 Método iterativo de Jacobi

La razón de haber estudiado primero la teoría de punto fijo para transformaciones afines, se debe a que todo sistema lineal se puede convertir en un problema de punto fijo equivalente, en el sentido de que toda solución del sistema lineal es un punto fijo de la transformación afín y viceversa.

► **Ejemplo 8.22** Considerar el sistema lineal ralo

$$\begin{array}{rcrcrcrcrcr} 10x_1 & - & x_2 & & & + & x_4 & & & = & 14 \\ 5x_1 & + & 40x_2 & & & & & + & 4x_5 & = & -43 \\ & & & + & 20x_3 & - & 4x_4 & & & = & 28 \\ & - & 2x_2 & & & - & 30x_4 & + & 2x_5 & = & -92 \\ 4x_1 & & & + & 2x_3 & & & - & 12x_5 & = & 32 \end{array} \qquad (8.30)$$

1. Encontrar una función afín, $g(\vec{x}) = T\vec{x} + \vec{c}$, en $\mathbb{R}^5$, cuyos puntos fijos coincidan con las soluciones del sistema lineal.
2. Probar, utilizando el teorema 8.3 con la norma matricial natural $\|\cdot\|^{(\infty)}$, que la función afín g tiene un único punto fijo $\vec{p}$ y que la sucesión de punto fijo (8.28) converge a $\vec{p}$ para todo valor inicial $\vec{x}_0$.
3. Tomando como valor inicial $\vec{x}_0 = \vec{0}_{\mathbb{R}^5}$, utilizar la cota para el error, relación (8.29), para calcular el menor número k de iteraciones necesarias con el fin de cometer un error menor a 10^{-3} al aproximar $\vec{p}$ con $\vec{x}_k$.
4. Con el valor k, encontrado en el inciso anterior, calcular las primeras k iteraciones con la relación de recurrencia (8.28). ¿Con cuál de ellas se garantiza un error menor a 10^{-3}?◄

Solución 1. Despejemos, en forma sucesiva, de cada una de las ecuaciones del sistema (8.30) a x_1, x_2, x_3, x_4 y x_5 para obtener el sistema equivalente;

$$\begin{aligned} x_1 &= \frac{1}{10}x_2 - \frac{1}{10}x_4 + \frac{7}{5} \\ x_2 &= -\frac{1}{8}x_1 - \frac{1}{10}x_5 - \frac{43}{40} \\ x_3 &= \frac{1}{5}x_4 + \frac{7}{5} \\ x_4 &= -\frac{1}{15}x_2 + \frac{1}{15}x_5 + \frac{46}{15} \\ x_5 &= \frac{1}{3}x_1 + \frac{1}{6}x_3 - \frac{8}{3} \end{aligned}$$

que en forma matricial se escribe

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{10} & 0 & -\frac{1}{10} & 0 \\ -\frac{1}{8} & 0 & 0 & 0 & -\frac{1}{10} \\ 0 & 0 & 0 & \frac{1}{5} & 0 \\ 0 & -\frac{1}{15} & 0 & 0 & \frac{1}{15} \\ \frac{1}{3} & 0 & \frac{1}{6} & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} + \begin{bmatrix} \frac{7}{5} \\ -\frac{43}{40} \\ \frac{7}{5} \\ \frac{46}{15} \\ -\frac{8}{3} \end{bmatrix}$$

Entonces, si

$$T = \begin{bmatrix} 0 & \frac{1}{10} & 0 & -\frac{1}{10} & 0 \\ -\frac{1}{8} & 0 & 0 & 0 & -\frac{1}{10} \\ 0 & 0 & 0 & \frac{1}{5} & 0 \\ 0 & -\frac{1}{15} & 0 & 0 & \frac{1}{15} \\ \frac{1}{3} & 0 & \frac{1}{6} & 0 & 0 \end{bmatrix}, \ \vec{c} = \begin{bmatrix} \frac{7}{5} \\ -\frac{43}{40} \\ \frac{7}{5} \\ \frac{46}{5} \\ -\frac{8}{3} \end{bmatrix} \text{ y } \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

toda solución del sistema (8.30) es punto fijo de la transformación afín

$$g(\vec{x}) = T\vec{x} + \vec{c}$$

y viceversa.

2. Puesto que

$$\begin{aligned} \|T\|^{(\infty)} &= \text{máx}\left(\frac{1}{10} + \frac{1}{10}, \frac{1}{8} + \frac{1}{10}, \frac{1}{5}, \frac{1}{15} + \frac{1}{15}, \frac{1}{3} + \frac{1}{6}\right) \\ &= \text{máx}\left(\frac{1}{5}, \frac{9}{40}, \frac{1}{5}, \frac{2}{15}, \frac{1}{2}\right) \\ &= \frac{1}{2} \end{aligned}$$

del teorema 8.3, la transformación g tiene un único punto fijo $\vec{p} \in \mathbb{R}^n$; que es, por tanto, la única solución del sistema (8.30); y la sucesión (8.28) converge a $\vec{p}$, respecto a la norma vectorial $\|\cdot\|_\infty$ para cualquier valor inicial $\vec{x}_0$.

3. En particular, por el inciso anterior, la sucesión (8.28) converge a $\vec{p}$ para el valor inicial $\vec{x}_0 = (0,0,0,0,0)$. Para acotar el error en la k-ésima iteración de punto fijo, dada por la relación (8.29) del corolario 8.29, necesitamos calcular la primera iteración

$$\begin{aligned} \vec{x}_1 &= g(\vec{x}_0) \\ &= T\vec{x}_0 + \vec{c} \\ &= \left(\frac{7}{5}, -\frac{43}{40}, \frac{7}{5}, \frac{46}{15}, -\frac{8}{3}\right). \end{aligned}$$

Por otra parte

$$\begin{aligned} \|\vec{x}_1 - \vec{x}_0\|_\infty &= \left\|\left(\frac{7}{5}, -\frac{43}{40}, \frac{7}{5}, \frac{46}{15}, -\frac{8}{3}\right)\right\|_\infty \\ &= \frac{46}{15}. \end{aligned}$$

Entonces, al aplicar (8.29), se tiene

$$\begin{aligned}\|\vec{x}_k-\vec{p}\|_\infty &\le \frac{(\|T\|^{(\infty)})^k}{1-\|T\|^{(\infty)}}\|\vec{x}_1-\vec{x}_0\|_\infty\\ &= \frac{(1/2)^k}{1-1/2}\|\vec{x}_1-\vec{x}_0\|_\infty\\ &= \frac{2\cdot 46}{15}\cdot\frac{1}{2^k}.\end{aligned}$$

Puesto que se quiere encontrar la menor iteración para un error inferior a 10^{-3}, basta encontrar el menor entero $k \geq 1$ tal que

$$\frac{2\cdot 46}{15}\cdot\frac{1}{2^k} < 10^{-3}.$$

Al resolver la precedente desigualdad para k, obtenemos

$$\begin{aligned}k &> \frac{\ln(96000/15)}{\ln(2)}\\ &\approx 12.6439.\end{aligned}$$

Luego, el menor número de iteraciones para el margen de error dado es

$$k = 13.$$

4. La tabla 8.1 contiene las primeras trece iteraciones realizadas en una computadora. Así que

$$x_{13} = \begin{bmatrix} 1.0000000001\\ -1.0000000000\\ 1.9999999999\\ 2.9999999999\\ -1.9999999999\end{bmatrix}$$

es la aproximación que produce un error inferior a 10^{-3}. De hecho, como el lector puede fácilmente verificar, la solución exacta del sistema (8.30) es

$$\vec{p} = \begin{bmatrix} 1\\ -1\\ 2\\ 3\\ -2\end{bmatrix}.$$

Evidentemente $\|\vec{x}_{13}-\vec{p}\|_\infty < 10^{-3}$. ✔

En la tabla 8.1 se observa claramente cómo las iteraciones $\vec{x}_k$ se van haciendo cada vez más próximas a $\vec{p}$ en la medida que k aumenta; esto es

$$\lim_{k\to\infty}\vec{x}_k = \vec{p}$$

respecto a la norma $\|\cdot\|_\infty$ de $\mathbb{R}^n$; pero, ya que este espacio tiene dimensión finita, en él todas las normas son equivalentes (cfr. teorema 4.22 de 4.2.4) y, por tanto, las iteraciones convergen a $\vec{p}$ respecto a cual-

x_0	x_1	x_2	x_3
0.0000000000	1.4000000000	0.9858333333	1.0056111111
0.0000000000	-1.0750000000	-0.9833333333	-1.0015625000
0.0000000000	1.4000000000	2.0133333333	1.9921111111
0.0000000000	3.0666666667	2.9605555556	3.0011111111
0.0000000000	-2.6666666667	-1.9666666667	-2.0025000000

x_4	x_5	x_6	x_7
0.9997326389	0.9999611111	0.9999910735	1.0000012066
-1.0004513889	-1.0000221354	-0.9999899306	-0.9999973796
2.0002222222	1.9999875000	2.0000134259	1.9999996007
2.9999375000	3.0000671296	2.9999980035	2.9999983256
-1.9994444444	-2.0000520833	-2.0000150463	-2.0000007378

x_8	x_9	x_10	x_11
1.0000004295	1.0000000147	0.9999999885	0.9999999978
-1.0000000770	-1.0000000873	-1.0000000106	-0.9999999983
1.9999996651	1.9999999552	2.0000000055	2.0000000023
2.9999997761	3.0000000275	3.0000000116	3.0000000005
-1.9999996644	-1.9999999127	-2.0000000026	-2.0000000029

x_12	x_13
1.0000000001	1.0000000001
-0.9999999994	-1.0000000000
2.0000000001	1.9999999999
2.9999999997	2.9999999999
-2.0000000004	-1.9999999999

Tabla 8-1 •

x_0	x_1	x_2	x_3
1000.00000000	-39.60000000	-121.51416667	-16.53744444
-4000.00000000	-626.07500000	-79.19166667	27.64427083
3000.00000000	-716.60000000	122.01333333	21.43655556
-3590.00000000	603.06666667	100.18277778	-0.67388889
5000.00000000	830.66666667	-135.30000000	-22.83583333

x_4	x_5	x_6	x_7
4.23181597	1.75744375	1.03154721	0.97366288
3.27576389	-1.14333811	-1.19016137	-1.01819595
1.26522222	1.34026528	1.90823796	2.01464196
-0.29867361	2.54118981	3.07320981	3.02217912
-4.60638889	-1.04519097	-1.85747454	-2.00477794

x_8	x_9	x_10	x_11
0.99596249	1.00028754	1.00018125	1.00001411
-0.99623007	-0.99886144	-0.99997529	-1.00003522
2.00443582	2.00017891	1.99986522	1.99997673
3.00089453	2.99932609	2.99988366	3.00000673
-2.00633871	-2.00060653	-1.99987434	-1.99996205

Tabla 8-2 •

quier norma en $\mathbb{R}^n$. La tabla 8.2 contiene las iteraciones de punto fijo, hasta $k = 11$, para resolver el ejemplo precedente, pero con un punto inicial más alejado de la solución; en ella puede observarse nuevamente cómo las iteraciones se van aproximando a la solución exacta aunque con menor rapidez.

En general, para un sistema lineal $A\vec{x} = \vec{b}$, $A = [a_{ij}]$, $\vec{x} = (x_1, x_2, \ldots, x_n)$, $\vec{b} = (b_1, b_2, \ldots b_n)$, podemos seguir la idea del ejemplo anterior para obtener una matriz T y un vector $\vec{c}$, de tal manera que las soluciones del sistema de ecuaciones y los puntos fijos de la función afín $g(\vec{x}) = T\vec{x} + \vec{c}$ coincidan, despejando cada una de las variables x_i de cada ecuación del sistema lineal. Es fácil comprobar que, haciendo eso,

$$T = I_n - B, \tag{8.31}$$

donde la fila i de la matriz B se obtiene multiplicando cada fila i de la matriz $A = [a_{ij}]$ por $1/a_{ii}$; mientras que al hacer lo propio con el vector $\vec{b}$ se obtiene al vector $\vec{c}$; donde hemos supuesto que los elementos en la diagonal de la matriz A son todos no nulos. El siguiente algoritmo contiene este método, llamado **método iterativo de Jacobi**, para aproximar la única solución de un sistema lineal, calculando el menor número de iteraciones para que el error que se cometa sea inferior a una cantidad que se fija de antemano. Utilizaremos la norma matricial natural $\|\cdot\|^{(\infty)}$ subordinada a la norma vectorial $\|\cdot\|_\infty$.

Método de Jacobi con cota para el error

Algoritmo 8.5

Entrada: *La matriz de coeficientes, $A = [a_{ij}] \in \mathfrak{M}_n(\mathbb{R})$, del sistema $A\vec{x} = \vec{b}$, donde $a_{ii} \neq 0$ para todo i, el término independiente $\vec{b} = [b_1 \ \cdots \ b_n]^t$; la aproximación inicial $\vec{x}_0$ y la cota para el margen de error máximo ε.*

Salida: *La primera iteración de punto fijo, $\vec{x}_k$, que aproxima a la solución exacta, $\vec{p}$, con*

$$\|\vec{x}_k - \vec{p}\| < \varepsilon$$

o el despliegue del texto "Posiblemente no hay convergencia" *en el caso de que $\|T\|^{(\infty)} \geq 1$.*

Paso 1. *Para cada $i = 1, 2, \ldots, n$, sean $F_i = (1/a_{ii})[\ a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}\ ]$ y*

$$T = I_n - B$$

donde B es la matriz cuyas filas son

$$F_i,\ i = 1, 2, \ldots, n;\ y\ \vec{c} = [\ b_1/a_{11} \quad b_2/a_{22} \quad \cdots \quad b_n/a_{nn}\ ].$$

Paso 2. *Se calcula la norma matricial natural, subordinada a la norma vectorial $\|\cdot\|_\infty$, de la matriz $T = [t_{ij}]$,*

$$\|T\|^{(\infty)} = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |t_{ij}|.$$

Si $\|T\|^{(\infty)} < 1$, se siguen los pasos 3 a 6; en caso contrario se despliega el mensaje "Posiblemente el sistema no tiene solución única".

Paso 3. *Se calculan la primera iteración de punto fijo,*

$$\vec{x}_1 = T\vec{x}_0 + \vec{c},$$

y

$$\|\vec{x}_1 - \vec{x}_0\|_\infty .$$

Paso 4. *Se calcula el menor entero k para el cual se cumple*

$$k > \frac{\ln\left(\varepsilon(1 - \|T\|^{(\infty)})/\|\vec{x}_1 - \vec{x}_0\|_\infty\right)}{\ln\left(\|T\|^{(\infty)}\right)}.$$

Paso 5. *Para cada* $j = 1, 2, \ldots, k$, *se calcula* $\vec{x}_j = T\vec{x}_{j-1} + \vec{c}$
Paso 6. *Se despliega la salida* $\vec{x}_k$.

Programa en MATLAB para el método de Jacobi con cota para el error

La figura 8-10 contiene el código en MATLAB de la función `Met_iter_Jacobi_1` para implementar el algoritmo 8.5. Al inicio del código se encuentra la descripción e instrucciones de uso; también se han intercalado entre las instrucciones algunos comentarios extras para ayudar a la comprensión del programa.

► **Ejemplo 8.23** Utilizar la función `Met_iter_jacobi_1` para aproximar, con el menor número de iteraciones posibles, la única solución del sitema lineal del ejemplo 8.22 con un error inferior al valor dado; tomando como punto inicial a $\vec{x}_0$.

1. 0.001, $\vec{x}_0 = [\ 0 \quad 0 \quad 0 \quad 0 \quad 0\ ]$.
2. 0.5×10^{-4}, $\vec{x}_0 = [\ 0 \quad 0 \quad 0 \quad 0 \quad 0\ ]$.
3. 0.0001, $\vec{x}_0 = [\ 25 \quad -10 \quad 12 \quad 8 \quad 90\ ]$.◄

```
function Met_iter_Jacobi_1(A,b,x0,epsilon)
%***Met_iter_Jacob_1(A,b,x0,epsilon)***
%Esta función calcula la menor iteración x_k con el método
%de Jacobi de la única solución p del sistema Ax=b
%con punto inicial x0; para un error ||x_k-p||_inf<epsilon
%si la norma matricial natural ||T||^(inf)<1;
%en caso de que ||T||^(inf)>=1, despliega un mensaje para señalar
%que posiblemente no hay convergencia.

n=size(A,1);
for i=1:n
   B(i,:)=(1/A(i,i))*A(i,:);
end
T=eye(n)-B;%%%matriz T
c=b./diag(A);%%vector c
N=norm(T,inf);%%%la norma N^(inf) de T.
if N<1
   x(:,1)=T*x0+c;%%%x1
   %%%%%%%k tiene que ser mayor a
   beta=log( ((1-N)/norm(x(:,1)-x0,inf))*epsilon)/log(N);
   %%%%% calculando k
   if (round(beta)-beta>=0)
   k=round(beta);
   else
   k=round(beta)+1;
   end
   %%%%Iteraciones
   for j=2:k
      x(:,j)=T*x(:,j-1)+c;
   end
   %%%%Imprimiendo x_k
   fprintf('\n\n x_%2.0f =\n\n',k);
   fprintf('  %14.10f\n', x(:,k));
   fprintf('\n')
   else
   %%%%%Mensaje si N>=1
   fprintf('\n Posiblemente no hay convergencia: ||T||^(inf)=%3.0f>=1\n\n',N)
end
```

Figura 8-10 • Función `Met_iter_Jacobi_1`.

Solución
```
EDU» M=[10 -1 0 1 0;5 40 0 0 4;0 0 20 -4 0; 0 -2 0 -30 2;4 0 20-12];
EDU» b=[14;-43;28;-92;32]; x0=[0 0 0 0 0]';
EDU»
```

1. `EDU» Met_iter_jacob1(M,b,x0,.001)`

```
x_13 =
        1.0000000001
       -1.0000000000
        1.9999999999
        2.9999999999
       -1.9999999999

EDU»
```

2. `EDU» Met_iter_jacob1(M,b,x0,.5*.0001)`

```
x_17 =
        1.0000000000
       -1.0000000000
        2.0000000000
        3.0000000000
       -2.0000000000

EDU»
```

3. `EDU» Met_iter_jacob1(M,b,[25 -10 12 8 90]',0.0001)`

```
x_21 =
        1.0000000000
       -1.0000000000
        2.0000000000
        3.0000000000
       -2.0000000000

EDU»  ✔
```

► **Ejemplo 8.24** Utilizar la función `Met_iter_jacobi1` para aproximar, con el menor número de iteraciones posibles, la única solución del sistema lineal

$$\begin{array}{rcrcrcr} x_1 & - & x_2 & + & 2x_3 & = & 4 \\ x_1 & - & 2x_2 & + & x_3 & = & 4 \\ 3x_1 & - & 3x_2 & + & 5x_3 & = & 11 \end{array}$$

con un error inferior a 0.0001; tomando como punto inicial a $\vec{x}_0 = \vec{0}_{\mathbb{R}^3}$.◄

Solución
```
EDU» E=[1 -1 2;1 -2 1;3 -3 5];
EDU» Met_iter_jacob1(E,[4 4 11]',[0 0 0]',.0001)
Posiblemente no hay convergencia:  ||T||^(inf)= 3>=1
EDU»  ✔
```

La razón de que en el algoritmo 8.5 y en su implementación en MATLAB se despliegue el mensaje *"Posiblemente no hay convergencia"* cuando $\|T\|^{(\infty)} \geq 1$, es debido a que cuando esto sucede la sucesión de punto fijo puede o no ser convergente. De no ser convergente, aun así es posible que el sistema lineal tenga solución única. Los siguientes dos ejemplos ilustran esta situación.

► **Ejemplo 8.25** Consideremos el sistema

$$\begin{array}{rcrcr} 2x & & & = & 2 \\ -4x & + & y & = & -1 \end{array}$$

Entonces, al despejar x, y de sendas ecuaciones, obtenemos

$$\begin{array}{lcl} x & = & 1 \\ \quad y & = & 4x-1 \end{array}$$

Luego, si

$$T = \begin{bmatrix} 0 & 0 \\ 4 & 0 \end{bmatrix} \text{ y } \vec{c} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

la única solución, $\vec{p} = (1,3)$, del sistema lineal es el único punto fijo de la transformación lineal $g(x) = T\vec{x} + \vec{c}$. Por otro lado,

$$\|T\|^{(\infty)} = 4.$$

y, si $\vec{x}_0 = (x_0, y_0)$,

$$\begin{aligned} \vec{x}_1 &= T\vec{x}_0 + \vec{c} \\ &= \begin{bmatrix} 0 & 0 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 4x_0 - 1 \end{bmatrix}, \\ \vec{x}_2 &= \begin{bmatrix} 0 & 0 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 4x_0 - 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ \vec{x}_3 &= \begin{bmatrix} 0 & 0 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 3 \end{bmatrix} \end{aligned}$$

etcétera,

$$\vec{x}_k = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \vec{p}.$$

Es decir,

$$\lim_{k\to\infty} \vec{x}_k = (1,3) = \vec{p}.$$

A pesar de que $\|T\|^{(\infty)} = 4 > 1$, la sucesión de punto fijo converge a la única solución del sistema lineal. ◀

▶ **Ejemplo 8.26** Ahora consideremos el sistema lineal

$$\begin{array}{rcrcl} 2x & - & 2y & = & 0 \\ -4x & + & y & = & 0 \end{array}.$$

Que tiene solución única $\vec{p} = (0,0)$. Entonces, dado que

$$x = y,$$
$$y = 4x$$

la transformación afín correspondiente resulta $g(\vec{x}) = T\vec{x}$, donde

$$T = \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix}.$$

Nuevamente

$$\|T\|^{(\infty)} = 4 > 1.$$

Sea $\vec{x}_0 = (a,b)$ cualquier valor inicial del espacio $\mathbb{R}^2$; entonces las primeras iteraciones de la sucesión de punto fijo están dadas por

$$\begin{aligned}
\vec{x}_1 &= \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \\
&= \begin{bmatrix} b \\ 4a \end{bmatrix}, \\
\vec{x}_2 &= \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} b \\ 4a \end{bmatrix} \\
&= \begin{bmatrix} 4a \\ 4b \end{bmatrix}, \\
\vec{x}_3 &= \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} 4a \\ 4b \end{bmatrix} \\
&= \begin{bmatrix} 4b \\ 4^2a \end{bmatrix}, \\
\vec{x}_4 &= \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} 4b \\ 4^2a \end{bmatrix} \\
&= \begin{bmatrix} 4^2a \\ 4^2b \end{bmatrix}, \\
\vec{x}_5 &= \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} 4^2a \\ 4^2b \end{bmatrix} \\
&= \begin{bmatrix} 4^2b \\ 4^3a \end{bmatrix}, \\
\vec{x}_6 &= \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} 4^2b \\ 4^3a \end{bmatrix} \\
&= \begin{bmatrix} 4^3a \\ 4^3b \end{bmatrix}.
\end{aligned}$$

De donde se desprende, por inducción, que

$$\vec{x}_{2k+1} = \begin{bmatrix} 4^k b \\ 4^{k+1}a \end{bmatrix}, \; k = 0,1,2,\ldots$$

y

$$\vec{x}_{2k} = \begin{bmatrix} 4^k a \\ 4^k b \end{bmatrix}, \; k = 1,2,3,\ldots$$

Luego, si $(a,b) \neq (0,0)$,

$$\lim_{k\to\infty} \vec{x}_k$$

no existe.◀

Sin embargo, conservando todavía la suposición[9] $\|T\|^{(\infty)} \geq 1$, si la sucesión de punto fijo (8.28) converge a un vector $\vec{p} \in \mathbb{R}^n$, entonces éste es un punto fijo de la transformación afín $g(\vec{x}) = T\vec{x} + \vec{c}$ y, por tanto, solución del sistema lineal $A\vec{x} = \vec{b}$. Efectivamente, por continuidad de la transformación afín g, se tiene

$$\begin{aligned} \vec{p} &= \lim_{k\to\infty} \vec{x}_k \\ &= \lim_{k\to\infty} g(\vec{x}_{k-1}) \\ &= g(\vec{p}). \end{aligned}$$

Pero, ya que la norma natural matricial de la matriz T es mayor a uno, $\vec{p}$ no necesariamente es solución única.

Tenemos entonces un problema al aplicar el algoritmo 8.5 cuando $\|T\|^{(\infty)} \geq 1$. Para soslayar esta dificultad, observemos que la sucesión $(\vec{x}_k)$ converge si y sólo si

$$\lim_{k\to\infty} \|\vec{x}_k - \vec{x}_{k-1}\|_\infty = 0.$$

(¡pruébelo!). Entonces, el valor $\|\vec{x}_k - \vec{x}_{k-1}\|_\infty$, en caso de convergencia, es cada vez más pequeño en la medida que k aumenta. Esto se observa cuando se aplica un ciclo iterativo convergente de punto fijo, notando que las cifras significativas entre iteraciones sucesivas comienzan a ser cada vez más cercanas (cfr. tabla 8.1). El algoritmo 8.6 es una versión del método de Jacobi para resolver en forma iterativa sistemas lineales, aun si la norma matricial natural es mayor o igual a la unidad. Aunque su carácter es empírico, lo expuesto previamente bien justifica su uso. Las modificaciones que se tienen que hacer al algortimo 8.6 son muy pocas y lo esencial consiste en un criterio de tolerancia, entre las diferencias de iteraciones sucesivas, para detener el proceso. Básicamente se calculan las iteraciones hasta que el valor $\|\vec{x}_k - \vec{x}_{k-1}\| < \delta$ aparece por primera vez, donde δ es el margen de tolerancia, que se fija previamente, para el cual se considera que la diferencia entre iteraciones sucesivas ya es suficientemente pequeña y, por tanto, la aproximación $\vec{x}_k$ es una buena estimación para resolver el problema. Si no se satisface la tolerancia para un determinado número (suficientemente grande) de iteraciones, entonces es plausible que no haya convergencia y el proceso debe detenerse; pues en caso de divergencia, como vimos antes, no se puede asegurar nada respecto a la existencia y unicidad de soluciones.

[9] Para fijar ideas, se ha estado trabajando respecto a la norma matricial natural $\|\cdot\|^{(\infty)}$, pero es evidente que las conclusiones son válidas para cualquier norma matricial natural subordinada a una norma vectorial.

Método de Jacobi con margen de tolerancia entre las diferencias de iteraciones sucesivas

Algoritmo 8.6

Entrada: *La matriz de coeficientes,* $A = [a_{ij}] \in \mathfrak{M}_n(\mathbb{R})$, *del sistema* $A\vec{x} = \vec{b}$, *donde* $a_{ii} \neq 0$ *para todo* i; *el término independiente* $\vec{b} = [b_1 \ \cdots \ b_n]^t$; *la aproximación inicial* $\vec{x}_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$; *la tolerancia* δ *para iteraciones sucesivas y el máximo número de iteraciones* N.

Salida: *La primera iteración de punto fijo,* $\vec{x}_k = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$, *que satisface el criterio*

$$\|\vec{x}_k - \vec{x}_{k-1}\| < \delta.$$

o el despliegue del texto "Máximo número de iteraciones excedido".

Paso 1. *Sea* $k = 1$.

Paso 2. *Mientras que* $k \leq N$ *siga los pasos 3 a 5.*

Paso 3. *Sean*

$$x_i^{(k)} = -\sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{a_{ij}}{a_{ii}} x_i^{(k-1)} + \frac{b_i}{a_{ii}}$$

y

$$\vec{x}_k = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$$

Paso 4. *Si* $\|\vec{x}_k - \vec{x}_{k-1}\| >= \delta$, *sea* $k = k+1$ *y se repiten los pasos 2 a 5.*

Paso 5. *Si* $\|\vec{x}_k - \vec{x}_{k-1}\| < \delta$, *se despliega la salida* $\vec{x}_k$.

Paso 6. *Si* $k = N$ *y* $\|\vec{x}_k - \vec{x}_{k-1}\| >= \delta$, *desplegar el mensaje* "Máximo número de iteraciones excedido".

❍ **Nota 8.5** El criterio de tolerancia en el algoritmo 8.6 para las diferencias entre iteraciones sucesivas,

$$|\vec{x}_k - \vec{x}_{k-1}| < \delta,$$

que aplica el error absoluto, se puede sustituir por el criterio

$$\frac{|\vec{x}_k - \vec{x}_{k-1}|}{|\vec{x}_k|} < \delta,$$

que aplica el error relativo, como sea conveniente.

► **Ejemplo 8.27** Resolver el ejemplo 8.22 (cfr. pág. 862), utilizando el algoritmo 8.6, para una tolerancia $\delta = 0.001$.◄

Solución Para resolver este problema utilizaremos los valores de las iteraciones $\vec{x}_k$ contenidos en la tabla 8.1. Tenemos entonces

$$\begin{aligned}
\|\vec{x}_1 - \vec{x}_0\|_\infty &= \left\| \begin{bmatrix} 1.4000000000 \\ -1.0750000000 \\ 1.4000000000 \\ 3.0666666667 \\ -2.6666666667 \end{bmatrix} \right\|_\infty \\
&= 3.0666666667 \\
&\geq 0.001 \,;
\end{aligned}$$

$$\|\vec{x}_2-\vec{x}_1\|_\infty = \left\| \begin{bmatrix} 0.9858333333 \\ -0.9833333333 \\ 2.0133333333 \\ 2.9605555556 \\ -1.9666666667 \end{bmatrix} - \begin{bmatrix} 1.4000000000 \\ -1.0750000000 \\ 1.4000000000 \\ 3.0666666667 \\ -2.6666666667 \end{bmatrix} \right\|_\infty$$

$$= \begin{bmatrix} -0.4141666667 \\ 0.0916666667 \\ 0.6133333333 \\ -0.1061111111 \\ 0.7000000000 \end{bmatrix}_\infty$$

$$= 0.7000000000$$

$$\geq 0.001\ ;$$

$$\|\vec{x}_3-\vec{x}_2\|_\infty = \left\| \begin{bmatrix} 1.0056111111 \\ -1.0015625000 \\ 1.9921111111 \\ 3.0011111111 \\ -2.0025000000 \end{bmatrix} - \begin{bmatrix} 0.9858333333 \\ -0.9833333333 \\ 2.0133333333 \\ 2.9605555556 \\ -1.9666666667 \end{bmatrix} \right\|_\infty$$

$$= \left\| \begin{bmatrix} 0.0197777778 \\ -0.0182291667 \\ -0.0212222222 \\ 0.0405555555 \\ -0.0358333333 \end{bmatrix} \right\|_\infty$$

$$= 0.0405555555$$

$$\geq 0.001\ ;$$

$$\|\vec{x}_4-\vec{x}_3\|_\infty = \left\| \begin{bmatrix} 0.9997326389 \\ -1.0004513889 \\ 2.0002222222 \\ 2.9999375000 \\ -1.9994444444 \end{bmatrix} - \begin{bmatrix} 1.0056111111 \\ -1.0015625000 \\ 1.9921111111 \\ 3.0011111111 \\ -2.0025000000 \end{bmatrix} \right\|_\infty$$

$$= \left\| \begin{bmatrix} -0.0058784722 \\ 0.0011111111 \\ 0.0081111111 \\ -0.0011736111 \\ 0.0030555556 \end{bmatrix} \right\|_\infty$$

$$= 0.0081111111$$

$$\geq 0.001\ ;$$

$$\|\vec{x}_5-\vec{x}_4\|_\infty = \left\| \begin{bmatrix} 0.9999611111 \\ -1.0000221354 \\ 1.9999875000 \\ 3.0000671296 \\ -2.0000520833 \end{bmatrix} - \begin{bmatrix} 0.9997326389 \\ -1.0004513889 \\ 2.0002222222 \\ 2.9999375000 \\ -1.9994444444 \end{bmatrix} \right\|_\infty$$

$$= \left\| \begin{bmatrix} 0.0002284722 \\ 0.0004292535 \\ -0.0002347222 \\ 0.0001296296 \\ -0.0006076389 \end{bmatrix} \right\|_\infty$$

$$= 0.0006076389$$

$$< 0.001\ .$$

Por lo que la primera iteración que satisface el margen de tolerancia $\|\vec{x}_k - \vec{x}_{k-1}\| < 0.001$ es

$$\vec{x}_5 = \begin{bmatrix} 0.9999611111 \\ -1.0000221354 \\ 1.9999875000 \\ 3.0000671296 \\ -2.0000520833 \end{bmatrix}. \checkmark$$

Programa en MATLAB para el método de Jacobi con margen de tolerancia entre iteraciones sucesivas

La figura 8-11 contiene la función `Met_iter_Jacobi_2` que implementa en MATLAB el algoritmo 8.6. Al inicio del mismo se encuentran las indicaciones para ejecutarlo. Se recomienda al lector que siga primero, con algún caso concreto, las instrucciones paso a paso para que por sí mismo pueda entender el funcionamiento de este programa.

```
function Met_iter_Jacobi_2(A,b,x0,delta,N)
%Met_iter_Jabobi_2(A,b,x0,delta,N)
%calcula la primera iteración x_k,
%para aproximar la solución del sistema
%Ax=b, con iteraciones de punto fijo
%por el método de Jacobi,
%para la cual ||x_k-x_k-1||< delta
%con punto inicial x0 y un máximo
%de iteraciones N. Si la tolerancia
%no se cumple a las N iteraciones
%se despliega un mensaje.
clear x;
n=size(A,1);
k=1;
while k<=N
   x(1,k)=-(1/A(1,1))*A(1,2:n)*x0(2:n)+b(1)/A(1,1);
for i=2:n
   x(i,k)=-(1/A(i,i))*(A(i,1:i-1)*x0(1:i-1)+A(i,i+1:n)*x0(i+1:n))+b(i)/A(i,i);
end
if norm(x(:,k)-x0,inf)>=delta
   x0=x(:,k);
   k=k+1;
else
   u=x(:,k);
   break
end
end
if k>N
fprintf('\n\n Número máximo de iteraciones excedido: N=%1.0f\n\n',N) else
   fprintf('\n\n x_%1.0f =\n\n',k);
   fprintf('  %14.10f\n', u);
fprintf('\n')
end
```

Figura 8-11 • Función `Met_iter_Jacobi_2`.

▶ **Ejemplo 8.28** Resolver el ejemplo 8.27 con la función `Met_iter_Jacobi_2` en MATLAB.◀

Solución EDU» F=[10 -1 0 1 0;5 40 0 0 4;
0 0 20 -4 0;0 -2 0 -30 2;
4 0 2 0 -12];

```
EDU» b=[14;-43;28;-92;32];x0=[0 0 0 0 0]';
EDU» Met_iter_Jacobi_2(F,b,x0,.001,20)
  x_5 =
       0.9999611111
      -1.0000221354
       1.9999875000
       3.0000671296
      -2.0000520833
EDU»  ✔
```

▶ **Ejemplo 8.29** Resolver el ejemplo anterior para una tolerancia $\delta = 0.000001$ con la función `Met_iter_Jacobi_2` en MATLAB.◀

Solución `EDU» Met_iter_Jacobi_2(F,b,x0,.000001,100)`

```
 x_9 =
      1.0000000147
     -1.0000000873
      1.9999999552
      3.0000000275
     -1.9999999127
EDU»  ✔
```

▶ **Ejemplo 8.30** Aplicar la función `Met_iter_Jacobi_2` al sistema lineal

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & & & = & -1 \\ & + & 2x_2 & + & x_3 & = & -1 \\ & + & 4x_2 & + & 3x_3 & = & -1 \end{array}$$

con una tolerancia $\delta = 0.00001$.◀

Solución Antes, observemos que en el sistema afín correspondiente, $g(\vec{x}) = T\vec{x} + \vec{c}$, al aplicar el método de Jacobi,

$$T = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & -1/2 \\ 0 & 4/3 & 0 \end{bmatrix};$$

entonces

$$\|T\|^{(\infty)} = 2 > 1.$$

```
EDU» A=[1 2 0;0 2 1;0 4 3];b=[-1 -1 -1]';
EDU» x0=[0 0 0]';
EDU» Met_iter_Jacobi_2(A,b,x0,.00001,20)
EDU» Número máximo de iteraciones excedido: N=20
EDU»
EDU» Met_iter_Jacobi_2(A,b,x0,.00001,100)
 x_55 =
     0.9999647981
    -0.9999911995
     0.9999765321
EDU»  ✔
```

La solución exacta del ejemplo anterior es $\vec{p} = (1, -1, 1)$, lo cual puede constatar fácilmente el lector.

Matrices diagonalmente dominantes y convergencia de las iteraciones del método de Jacobi

Hasta aquí hemos visto solamente un criterio para identificar la convergencia de las iteraciones de punto fijo del método de Jacobi; comprobando directamente que $\|T\|^{(\infty)} < 1$ para la matriz T de la transformación afín $g(\vec{x}) = T\vec{x} + \vec{c}$ que corresponde al sistema lineal $A\vec{x} = \vec{b}$. Existe una clase de matrices, llamadas diagonalmente dominantes, para las cuales la iteración de Jacobi converge independientemente del valor inicial $\vec{x}_0$. De hecho, como veremos más adelante, ese tipo de matrices son invertibles y, por tanto, el sistema lineal $A\vec{x} = \vec{b}$ tiene solución única para todo $\vec{b}$; más aún, el método de Gauss (algoritmo 8.1) se puede aplicar, para resolver el sistema, sin hacer cambios de renglones y existe estabilidad respecto al crecimiento en errores de redondeo.

Definición 8.4 *Sea A una matriz cuadrada de orden n. Se dice que $A = [a_{ij}]$ es* ***diagonalmente dominante****, si*

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$$

para cada $i = 1, 2, \ldots$

Es decir, una matriz es diagonalmente dominante si el valor absoluto de cada elemento de la diagonal es mayor que la suma de los valores absolutos de las demás componentes de su renglón.

► **Ejemplo 8.31** Dado que para la matriz

$$A = \begin{bmatrix} 4 & -1 & 2 \\ 1 & 3 & 1 \\ -2 & 1 & 5 \end{bmatrix}$$

se cumple $|4| > |-1| + |2|$, $|3| > |1| + |1|$ y $|5| > |-2| + |1|$, la matriz A es diagonalmente dominante. ◄

Vimos, en la discusión ulterior al ejemplo 8.22, que si A es una matriz cuadrada de orden n cuyos elementos en la diagonal son no nulos, entonces las iteraciones de punto fijo del método de Jacobi se aplican a la transformación afín $g(\vec{x}) = T\vec{x} + \vec{c}$, donde $T = I_n - B$ y B es la matriz que se obtiene al multiplicar cada fila i de la matriz A por el escalar $1/a_{ii}$; esto es,

$$T = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \cdots & -\frac{a_{1(n-1)}}{a_{11}} & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} & \vdots & -\frac{a_{2(n-1)}}{a_{11}} & -\frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ & & & \ddots & & \\ -\frac{a_{(n-1)1}}{a_{(n-1)(n-1)}} & -\frac{a_{(n-1)2}}{a_{(n-1)(n-1)}} & \cdots & \cdots & 0 & -\frac{a_{(n-1)n}}{a_{(n-1)(n-1)}} \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \cdots & \cdots & -\frac{a_{n(n-1)}}{a_{nn}} & 0 \end{bmatrix} \tag{8.32}$$

Por tanto, si suponemos que A es una matriz diagonalmente dominante, entonces

$$\begin{aligned} \|T\|^{(\infty)} &= \max_{1 \leq i \leq j} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{|a_{ij}|}{|a_{ii}|} = \max_{1 \leq i \leq n} \frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \\ &< 1. \end{aligned}$$

Del teorema 8.3 se concluye que la iteración de punto fijo de Jacobi converge a la única solución del sistema $A\vec{x}=\vec{b}$ para cualquier punto inicial $\vec{x}_0$. Hemos probado así el siguiente teorema.

Teorema 8.4 *Si A es una matriz cuadrada de orden n que es diagonalmente dominante, entonces la sucesión de punto fijo que resulta del método de Jacobi converge a la única solución del sistema lineal $A\vec{x}=\vec{b}$ para cualquier vector inicial $\vec{x}_0$.*

8.3.3 Planteamiento general para un método iterativo

En el apartado anterior vimos que dado un sistema lineal de ecuaciones $A\vec{x}=\vec{b}$, $A=[a_{ij}]$, $\vec{b}=[b_i]^t$, es posible encontrar una matriz T y un vector $\vec{c}$ de tal manera que los puntos fijos de la transformación afín $g(\vec{x})=T\vec{x}+\vec{c}$ coinciden con las soluciones del sistema y, entonces, a partir de las iteraciones de punto fijo, aproximar la solución. Para el método de Jacobi, la matriz T está dada por (8.32). Pero si $D=\operatorname{diag}(a_{11},a_{22},\ldots,a_{nn})$; es decir, D es la matriz diagonal cuadrada cuya diagonal son los elementos a_{ii} de A (cfr. definición 1.4, pág 8), entonces $D^{-1}=\operatorname{diag}(1/a_{11},1/a_{22},\ldots,1/a_{nn})$ —suponiendo que todos los a_{ii} son distintos de cero— y, por tanto,

$$\begin{aligned} T &= I_n - D^{-1}A,\\ \vec{c} &= D^{-1}\vec{b}. \end{aligned}$$

En general, podemos elegir, de manera análoga, cualquier matriz Q de orden n que sea invertible y hacer

$$\left.\begin{aligned} T_Q &= I_n - Q^{-1}A\\ \vec{c} &= Q^{-1}\vec{b} \end{aligned}\right\} \tag{8.33}$$

para obtener la transformación afín

$$g_Q(\vec{x}) = T_Q\vec{x}+\vec{c} \tag{8.34}$$

cuyo conjunto de puntos fijos coincide con las soluciones del sistema lineal $A\vec{x}=\vec{b}$. En efecto:

$$\begin{aligned} A\vec{p}=\vec{b} &\Leftrightarrow (Q^{-1}A)\vec{p}=Q^{-1}\vec{b}\\ &\Leftrightarrow \vec{p}-(Q^{-1}A)\vec{p}=\vec{p}-Q^{-1}\vec{b}\\ &\Leftrightarrow (I_n-Q^{-1}A)\vec{p}+Q^{-1}\vec{b}=\vec{p}\\ &\Leftrightarrow T_Q\vec{p}+\vec{c}=\vec{p}\\ &\Leftrightarrow g_Q(\vec{p})=\vec{p}. \end{aligned}$$

Así, podemos aplicar la teoría del punto fijo a la transformación afín (8.34) para iterar la sucesión

$$\vec{x}_k = \begin{cases} \vec{x}_0, & \text{si } k=0\\ g_Q(\vec{x}_{k-1}) & \text{si } k=1,2,\ldots \end{cases} \tag{8.35}$$

para cualquier matriz Q que sea invertible con el fin de aproximar la solución del sistema lineal. Diremos entonces que la relación de recurrencia (8.35) es **un método iterativo** para aproximar la solución del sistema $A\vec{x} = \vec{b}$ por medio de la transformación afín g_Q. La convergencia quedará asegurada si $\|T_Q\| < 1$; esto es, si

$$\left\|I_n - Q^{-1}A\right\| < 1 \tag{8.36}$$

para alguna norma matricial natural.

Sin embargo, debido a la dificultad que entraña el cálculo de la inversa de una matriz, la condición (8.36) es poco viable de aplicar en la práctica; de hecho, por la misma razón, el método iterativo $\vec{x}_k = T_Q\vec{x}_{k-1} + \vec{c}$ utilizando directamente la matriz T_Q es poco probable de implementar y, en lugar de ello, se buscan relaciones de recurrencia a partir del propio sistema lineal como hicimos en el método de Jacobi del algorimo 8.6. A pesar de ello, la condición suficiente de convergencia (8.36) se puede deducir a partir de condiciones más simples que tienen que ver con la matriz de coeficientes A, la cual tenemos a la mano para analizar sin tener que calcular la matriz T_Q para un método específico. Para poder establecer esas condiciones, requerimos de los siguientes conceptos.

Definición 8.5 **(Radio espectral)** *Sea A una matriz cuadrada de orden n. Se llama* ***radio espectral*** *de la matriz A al número denotado y definido por*

$$\rho(A) = \text{máx}\,\{|\lambda| : \lambda \text{ es valor propio de } A\}\,.$$

Es decir, el radio espectral es la máxima magnitud de los valores propios de una matriz. El radio tiene dos propiedades fundamentales que hacemos patentes a continuación.

Proposición 8.4 *Sean A una matriz cuadrada de orden n y $\|\cdot\|$ cualquier norma matricial natural. Entonces*

$$\rho(A) \leq \|A\| \tag{8.37}$$

DEMOSTRACIÓN ■ Sea $\|\cdot\|$ una norma matricial natural en $\mathfrak{M}_n(\mathbb{R})$ que proviene de una norma vectorial $\|\cdot\|$. Sea λ cualquier valor propio de A y $\vec{u}$ un vector propio correspondiente con $\|\vec{u}\| = 1$; entonces, por 8.20 de la proposición 8.2 (pág. 857)

$$\begin{aligned} |\lambda| &= |\lambda|\,\|\vec{u}\| \\ &= \|\lambda\vec{u}\| \\ &= \|A\vec{u}\| \\ &\leq \|A\|\,\|\vec{u}\| \\ &= \|A\|, \end{aligned}$$

y, en consecuencia, se tiene (8.37). ■

Recordemos[10] que si S es un conjunto no vacío de números reales, un número real m es una **cota inferior** del conjunto S si $m \leq s$ para todo $s \in S$. Si S tiene una cota inferior se dice que el conjunto S

[10] Cfr. la discusión que precede al teorema 6.11 del apartado 6.8.2.

está **acotado inferiormente**. Un principio fundamental de las matemáticas establece que todo conjunto no vacío de números reales S que está acotado inferiormente posee una máxima cota inferior; es decir, existe un número real α, cota inferior de S, tal que $m \leq \alpha$ para toda cota inferior m de S. Se acostumbra entonces utilizar la notación

$$\alpha = \text{ínf}\, S$$

o

$$\alpha = \text{ínf}_{s \in S} s$$

y llamar a α el **ínfimo** de S. Es obvio que α no necesariamente pertenece a S; si $\alpha \in S$, α se llama entonces el mínimo de S. Ambos, el ínfimo y el mínimo, son únicos (¡demuéstrelo!).

Se puede probar que si $\mathfrak{N}$ es el conjunto de todas las normas matriciales naturales en el espacio $\mathfrak{M}_n(\mathbb{R})$, entonces[11]

$$\rho(A) = \text{ínf}_{\|\cdot\| \in \mathfrak{N}} \|A\| \tag{8.38}$$

Terminamos este breve, pero esencial apartado, con un teorema que plantea una condición para la convergencia del método iterativo (8.35) aplicando el concepto de radio espectral.

Teorema 8.5 *Sean A y Q un par de matrices cuadradas de orden n, $I_n \in \mathfrak{M}_n(\mathbb{R})$ la matriz identidad y $\vec{b} \in \mathbb{R}^n$. Si la matriz cuadrada Q es invertible, $T_Q = (I_n - Q^{-1}A)$, $\vec{c} = Q^{-1}\vec{b}$ y*

$$\rho(T_Q) = \rho(I_n - Q^{-1}A) < 1$$

entonces el método iterativo

$$\vec{x}_k = \begin{cases} \vec{x}_0, & \text{si } k = 0 \\ T_Q\vec{x}_{k-1} + \vec{c}, & \text{si } k = 1, 2, \ldots \end{cases}$$

converge a la única solución del sistema lineal $A\vec{x} = \vec{b}$ para cualquier vector inicial $\vec{x}_0$.

DEMOSTRACIÓN ■ Puesto que $\rho(T_Q) < 1$, entonces, por 8.38, existe una norma matricial natural $\|\cdot\|$ tal que

$$\rho(T_Q) \leq \|T_Q\| < 1$$

(pues en caso contrario 1 sería una cota inferior mayor que el radio espectral). Ya que $\|T_Q\| < 1$, por el teorema del punto fijo para transformaciones afines (teorema 8.3, pág. 861), el método iterativo converge a la única solución del sistema lineal. ■

[11] En realidad, la notación debería ser $\rho(A) = \text{ínf}\{\|A\| : \|\cdot\| \text{ es una norma matricial natural}\}$; pero la notación que hemos utilizado es compacta y fácil de comprender.

8.3.4 Método iterativo de Richardson

Un caso particularmente sencillo de la construcción (8.33) es elegir $Q = I_n$, con lo que

$$T = I_n - A,\ \vec{c} = \vec{b}$$

y se obtiene la transformación afín correspondiente

$$g(\vec{x}) = (I_n - A)\vec{x} + \vec{b}.$$

Entonces, la sucesión de punto fijo para la función g se reduce a

$$\begin{aligned} \vec{x}_k &= g(\vec{x}_{k-1}) \\ &= T\vec{x}_{k-1} + \vec{b} \end{aligned}$$

o

$$\vec{x}_k = \vec{x}_{k-1} - A\vec{x}_{k-1} + \vec{b}$$

De esta manera, si $\vec{x}_k = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$, con $\vec{x}_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$ la aproximación inicial,

$$x_i^{(k)} = x_i^{(k-1)} + b_i - \sum_{j=1}^{n} a_{ij} x_j^{(k-1)} \tag{8.39}$$

es la sucesión de iteraciones para aproximar la solución del sistema lineal. La relación de recurrencia (8.39) es el llamado **método iterativo de Richardson.** La motivación algebraica del método de Richardson es muy sencilla; pues ya que

$$A\vec{x} = \vec{b}$$

equivale a

$$\vec{x} + A\vec{x} = \vec{b} + \vec{x}$$

entonces,

$$\begin{aligned} \vec{x} &= \vec{x} - A\vec{x} + \vec{b} \\ &= (I_n - A)\vec{x} + \vec{b}. \end{aligned}$$

Ésta es la razón de la elección $Q = I_n$.

▶ **Ejemplo 8.32** Para el sistema lineal

$$\begin{array}{rcrcrcrcrcr} x_1 & & & & & + & \frac{1}{4}x_4 & + & \frac{1}{8}x_5 & = & 27 \\ & + & x_2 & & & + & \frac{1}{3}x_4 & + & \frac{1}{6}x_3 & = & -20 \\ \frac{1}{8}x_1 & - & \frac{1}{4}x_2 & + & x_3 & & & & & = & 57 \\ & + & \frac{1}{2}x_2 & & & + & x_4 & & & = & 12 \\ \frac{1}{3}x_1 & & & + & \frac{1}{3}x_3 & & & + & x_5 & = & 0 \end{array} \tag{8.40}$$

tenemos que

$$A=\begin{bmatrix}1 & 0 & 0 & \frac{1}{4} & \frac{1}{8}\\ 0 & 1 & 0 & \frac{1}{3} & \frac{1}{6}\\ \frac{1}{8} & -\frac{1}{4} & 1 & 0 & 0\\ 0 & \frac{1}{2} & 0 & 1 & 0\\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & 1\end{bmatrix} \quad \text{y} \quad \vec{b}=\begin{bmatrix}27\\-20\\57\\12\\0\end{bmatrix}.$$

Entonces,

$$\begin{aligned}T &= I_n - A\\ &= \begin{bmatrix}0 & 0 & 0 & -\frac{1}{4} & -\frac{1}{8}\\ 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{6}\\ -\frac{1}{8} & \frac{1}{4} & 0 & 0 & 0\\ 0 & -\frac{1}{2} & 0 & 0 & 0\\ -\frac{1}{3} & 0 & -\frac{1}{3} & 0 & 0\end{bmatrix}.\end{aligned}$$

Si $\vec{x}_0=(0,0,0,0,0)$, las primeras tres iteraciones del método de Richardson producen:

$$\begin{aligned}\vec{x}_1 &= T\vec{x}_0+\vec{b}\\ &= \begin{bmatrix}27\\-20\\57\\12\\0\end{bmatrix},\\ \vec{x}_2 &= T\vec{x}_1+\vec{b}\\ &= \begin{bmatrix}0 & 0 & 0 & -\frac{1}{4} & -\frac{1}{8}\\ 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{6}\\ -\frac{1}{8} & \frac{1}{4} & 0 & 0 & 0\\ 0 & -\frac{1}{2} & 0 & 0 & 0\\ -\frac{1}{3} & 0 & -\frac{1}{3} & 0 & 0\end{bmatrix}\begin{bmatrix}27\\-20\\57\\12\\0\end{bmatrix}+\begin{bmatrix}27\\-20\\57\\12\\0\end{bmatrix}\\ &= \begin{bmatrix}24.0\\-24.0\\48.625\\22.0\\-28.0\end{bmatrix} \quad \text{y}\\ \vec{x}_3 &= T\vec{x}_2+\vec{b}\\ &= \begin{bmatrix}0 & 0 & 0 & -\frac{1}{4} & -\frac{1}{8}\\ 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{6}\\ -\frac{1}{8} & \frac{1}{4} & 0 & 0 & 0\\ 0 & -\frac{1}{2} & 0 & 0 & 0\\ -\frac{1}{3} & 0 & -\frac{1}{3} & 0 & 0\end{bmatrix}\begin{bmatrix}24.0\\-24.0\\48.625\\22.0\\-28.0\end{bmatrix}+\begin{bmatrix}27\\-20\\57\\12\\0\end{bmatrix}\\ &= \begin{bmatrix}25.0\\-22.66666667\\48.0\\24.0\\-24.20833333\end{bmatrix}. \blacktriangleleft\end{aligned}$$

Las iteraciones de punto fijo para el método convergen si $\|T\| = \|I_n - A\| < 1$ para alguna norma matricial natural; y (8.29) del corolario 8.2 es una cota para el error con la k-ésima iteración.

▶ **Ejemplo 8.33** Encontrar el menor número de iteraciones, k, que produce el método iterativo de Richardson para garantizar un error máximo inferior a 0.01 al aproximar la solución del sistema (8.40) del ejemplo anterior. Trabajar el problema respecto a la norma matricial natural $\|\cdot\|^{(\infty)}$.◀

Solución En el ejemplo 8.32 encontramos

$$T = I_5 - A = \begin{bmatrix} 0 & 0 & 0 & -\frac{1}{4} & -\frac{1}{8} \\ 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{8} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 \\ -\frac{1}{3} & 0 & -\frac{1}{3} & 0 & 0 \end{bmatrix}.$$

Entonces, ya que

$$\begin{aligned} \|T\|^{(\infty)} &= \text{máx}\left(\frac{1}{4}+\frac{1}{8}, \frac{1}{3}+\frac{1}{6}, \frac{1}{8}+\frac{1}{4}, \frac{1}{2}, \frac{1}{3}+\frac{1}{3}\right) \\ &= \frac{2}{3} < 1 \end{aligned}$$

la sucesión de punto fijo del método de Richardson converge a la solución exacta, $\vec{p}$, del sistema (8.40), para cualquier valor inicial $\vec{x}_0$; además, por (8.29) del corolario 8.2,

$$\|\vec{x}_k - \vec{p}\|_\infty \le \frac{\left(\|T\|^{(\infty)}\right)^k}{1-\|T\|^{(\infty)}}\|\vec{x}_1 - \vec{x}_0\|_\infty .$$

Tomemos $\vec{x}_0 = (0,0,0,0,0)$ como punto inicial, entonces $\vec{x}_1 = \vec{b} = (27,-20,57,12,0)$; por tanto,

$$\begin{aligned} \|\vec{x}_k - \vec{p}\|_\infty &\le \frac{\left(\frac{2}{3}\right)^k}{1-\frac{2}{3}}\|\vec{x}_1 - \vec{x}_0\|_\infty \\ &= \frac{\left(\frac{2}{3}\right)^k}{1-\frac{2}{3}}57. \end{aligned}$$

Puesto que se necesita $\|\vec{x}_k - \vec{p}\|_\infty < 0.01$ para el menor entero k, entonces basta tomar el primer entero k que satisface

$$k > \frac{\ln(3\times 57\times 10^2)}{\ln(3/2)} \approx 24.03864981\,.$$

Así, se requieren de $k = 25$ iteraciones. La tabla 8.3, contiene las 25 primeras iteraciones realizadas en una computadora. El valor

$$\vec{x}_{25} = \begin{bmatrix} 24.00000005 \\ -23.99999994 \\ 48.00000002 \\ 23.99999995 \\ -24.00000004 \end{bmatrix}$$

satisface $\|\vec{x}_k - \vec{p}\|_\infty < 0.01$. ✔

```
     x_0          x_1          x_2          x_3          x_4
  0.00000000  27.00000000  24.00000000  25.00000000  24.02604167
  0.00000000 -20.00000000 -24.00000000 -22.66666667 -23.96527778
  0.00000000  57.00000000  48.62500000  48.00000000  48.20833333
  0.00000000  12.00000000  22.00000000  24.00000000  23.33333333
  0.00000000   0.00000000 -28.00000000 -24.20833333 -24.33333333

     x_5          x_6          x_7          x_8          x_9
 24.20833333  24.01410590  24.04362883  24.00474718  24.00921179
-23.72222222 -23.98119213 -23.94182822 -23.99367043 -23.98771762
 48.00542535  48.04340278  48.00293873  48.00908934  48.00098900
 23.98263889  23.86111111  23.99059606  23.97091411  23.99683521
-24.07812500 -24.07125289 -24.01916956 -24.01552252 -24.00461217

     x_10         x_11         x_12         x_13         x_14
 24.00136772  24.00196033  24.00036490  24.00042027  24.00009304
-23.99817638 -23.99738623 -23.99951346 -23.99943963 -23.99987595
 48.00191912  48.00028494  48.00040840  48.00007602  48.00008756
 23.99385881  23.99908819  23.99869311  23.99975673  23.99971982
-24.00340026 -24.00109561 -24.00074842 -24.00025777 -24.00016543

     x_15         x_16         x_17         x_18         x_19
 24.00009072  24.00002303  24.00001971  24.00000559  24.00000431
-23.99987903 -23.99996929 -23.99997372 -23.99999255 -23.99999426
 48.00001938  48.00001890  48.00000480  48.00000411  48.00000116
 23.99993797  23.99993952  23.99998465  23.99998686  23.99999628
-24.00006020 -24.00003670 -24.00001398 -24.00000817 -24.00000323

     x_20         x_21         x_22         x_23         x_24
 24.00000133  24.00000095  24.00000032  24.00000021  24.00000007
-23.99999822 -23.99999874 -23.99999958 -23.99999972 -23.99999990
 48.00000090  48.00000028  48.00000020  48.00000007  48.00000004
 23.99999713  23.99999911  23.99999937  23.99999979  23.99999986
-24.00000182 -24.00000074 -24.00000041 -24.00000017 -24.00000009

     x_25
 24.00000005
-23.99999994
 48.00000002
 23.99999995
-24.00000004
```

Tabla 8-3 •

Resumimos el procedimiento precedente en el algoritmo 8.7.

Método iterativo de Richardson con cota para el error

Algoritmo 8.7

Entrada: *La matriz de coeficientes,* $A = [a_{ij}] \in \mathfrak{M}_n(\mathbb{R})$*, del sistema* $A\vec{x} = \vec{b}$*; el término independiente* $\vec{b} = [b_1 \ \cdots \ b_n]^t$*; la aproximación inicial* $\vec{x}_0$ *y la cota para el margen de error máximo* ε*.*

Salida: *La primera iteración de punto fijo $\vec{x}_k$ que aproxima a la solución exacta $\vec{p}$, con*

$$\|\vec{x}_k - \vec{p}\| < \varepsilon.$$

o el despliegue del texto "Posiblemente no hay convergencia" *en el caso de que* $\|T\|^{(\infty)} \geq 1$.

Paso 1. *Se calcula la matriz*

$$T = I_n - A$$

Paso 2. *Se calcula la norma matricial natural, subordinada a la norma vectorial $\|\cdot\|_\infty$, de la matriz $T = [t_{ij}]$,*

$$\|T\|^{(\infty)} = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |t_{ij}|.$$

Si $\|T\|^{(\infty)} < 1$, se siguen los pasos 3 a 6; en caso contrario se despliega el mensaje "Posiblemente no hay convergencia".

Paso 3. *Se calculan la primera iteración de punto fijo,*

$$\vec{x}_1 = T\vec{x}_0 + \vec{c},$$

y

$$\|\vec{x}_1 - \vec{x}_0\|_\infty .$$

Paso 4. *Se calcula el menor entero k para el cual se cumple*

$$k > \frac{\ln\left(\varepsilon(1 - \|T\|^{(\infty)})/\,\|\vec{x}_1 - \vec{x}_0\|_\infty\right)}{\ln\left(\|T\|^{(\infty)}\right)}.$$

Paso 5. *Para cada $j = 1, 2, \ldots, k$, se calcula*

$$\vec{x}_j = T\vec{x}_{j-1} + \vec{b}$$

Paso 6. *Se despliega la salida*

$$\vec{x}_k.$$

Programa en MATLAB para el método de Richardson con cota para el error

La figura 8-12 contiene el código en MATLAB de la función `Met_iter_Richardson_1` para implementar el algoritmo 8.7. Al inicio del código se encuentra la descripción e instrucciones de uso.

▶ **Ejemplo 8.34** Resolver el ejemplo 8.33 aplicando la función `Met_iter_Richardson_1`.◀

Solución

```
EDU» A=[1 0 0 1/4 1/8;0 1 0 1/3 1/6;1/8 -1/4 1 0 0;
0 1/2 0 1 0;1/3 0 1/3 0 1];
EDU» b=[27 -20 57 12 0]';x0=[0 0 0 0 0]';
EDU» Met_iter_Richardson_1(A,b,x0,.01,40)
 x_25 =
    24.0000000462
   -23.9999999384
    48.0000000154
    23.9999999507
   -24.0000000391
EDU»
```

✔

```
function Met_iter_Richardson_1(A,b,x0,epsilon)
%***Met_iter_Richardson_1(A,b,x0,epsilon)***
%Esta función calcula la menor iteración x_k con el método
%de Richardson para aproximar la única solución p del sistema Ax=b
%con punto inicial x0; para un error ||x_k-p||_inf<epsilon
%si la norma matricial natural ||T||^(inf)<1;
%en caso de que ||T||^(inf)>=1, despliega un mensaje para señalar
%que posiblemente no hay convergencia.
clear x;
n=size(A,1);
T=eye(n)-A;%%%matriz T de la transformación afín.
N=norm(T,inf);%%%la norma N^(infinto) de T.
   if N<1
      x(:,1)=T*x0+b;%%%x1
      %%%%%%%k tiene que ser mayor a:
      beta=log( ((1-N)/norm(x(:,1)-x0,inf))*epsilon)/log(N);
      %%%%%Calculando k
      if (round(beta)-beta>=0)
      k=round(beta);
      else
      k=round(beta)+1;
      end
      %%%%Iteraciones
      for j=2:k
         x(:,j)=T*x(:,j-1)+b;
      end
%%%Impresión de x_k
fprintf('\n\n x_%2.0f =\n\n',k);
fprintf('  %14.10f\n', x(:,k));
fprintf('\n')
else
%%%%%Mensaje si N>=1
fprintf('\n Posiblemente no hay convergencia: ||T||^(inf)=%3.0f>=1\n\n',N)
end
```

Figura 8-12 • Función Met_iter_Richardson_1.

Obviamente, como en el caso del método de Jacobi, si $\|T\| = \|I_n - A\| \geq 1$, entonces es posible que las iteraciones del método de Richardson converjan o no. El siguiente procedimiento es la versión del algoritmo 8.6 para el método de Richardson.

Método de Richardson con margen de tolerancia para las diferencias de iteraciones sucesivas

Algoritmo 8.8

Entrada: *La matriz de coeficientes,* $A = [a_{ij}] \in \mathfrak{M}_n(\mathbb{R})$, *del sistema* $A\vec{x} = \vec{b}$; *el término independiente* $\vec{b} = [b_1 \cdots b_n]^t$; *la aproximación inicial* $\vec{x}_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$; *la tolerancia* δ *para la diferencia de iteraciones sucesivas y el máximo número de iteraciones* N.

Salida: *La primera iteración de punto fijo,* $\vec{x}_k = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$, *que satisface el criterio*

$$\|\vec{x}_k - \vec{x}_{k-1}\| < \delta$$

o el despliegue del texto "Máximo número de iteraciones excedido".

Paso 1. *Sea* $k = 1$.

Paso 2. *Mientras que* $k \leq N$ *siga los pasos 3 a 5.*

Paso 3. *Sean*

$$x_i^{(k)} = x_i^{(k-1)} + b_i - \sum_{j=1}^{n} a_{ij} x_j^{(k-1)}$$

y

$$\vec{x}_k = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$$

Paso 4. *Si* $\|\vec{x}_k - \vec{x}_{k-1}\| >= \delta$, *sea* $k = k+1$ *y se repiten los pasos 2 a 5.*

Paso 5. *Si* $\|\vec{x}_k - \vec{x}_{k-1}\| < \delta$, *se despliega la salida* $\vec{x}_k$.

Paso 6. *Si* $k = N$ *y* $\|\vec{x}_k - \vec{x}_{k-1}\| >= \delta$, *desplegar el mensaje* "Máximo número de iteraciones excedido".

Programa en MATLAB para el algoritmo 8.8

La figura 8-13 contiene la función `Met_iter_Richardson_2` para implementar el algoritmo 8.8 en el ambiente MATLAB. Las instrucciones están contenidas al inicio del código, que básicamente es el programa `Met_iter_Jacobi_2` (cfr. figura 8-11, pág. 874) con las adecuaciones pertinentes al método de Richardson.

► **Ejemplo 8.35** Resolver el ejemplo 8.33 aplicando la función `Met_iter_Richardson_2` para una tolerancia $\delta = 0.001$.◄

Solución

```
EDU» A=[1 0 0 1/4 1/8;0 1 0 1/3 1/6;1/8 -1/4 1 0 0;
0 1/2 0 1 0;1/3 0 1/3 0 1];
EDU» b=[27 -20 57 12 0]'; x0=[0 0 0 0 0];
EDU» Met_iter_Richardson_2(A,b,x0,.001,20)
 x_14 =
    24.0000930386
   -23.9998759486
    48.0000875572
    23.9997198168
   -24.0001654322
EDU»
```

✔

```
function Met_iter_Richardson_2(A,b,x0,delta,N)
%***Met_iter_Richardson_2(A,b,x0,delta,N)***
"%calcula la primera iteración x_k,"
%para aproximar la solución del sistema
%Ax=b, con iteraciones de punto fijo
%por el método de Richardson,
%para la cual ||x_k-x_k-1||< delta
%con punto inicial x0 y un máximo
%de iteraciones N. Si la tolerancia
%no se cumple a las N iteraciones
%se despliega un mensaje.
clear x;
n=size(A,1);
k=1;
while k<=N
 for i=1:n
   x(i,k)=x0(i)+b(i)-A(i,:)*x0;
 end if norm(x(:,k)-x0,inf)>=delta
   x0=x(:,k);
   k=k+1;
else
   u=x(:,k);
   break
end end if k>N

fprintf('\n\nNúmero máximo de iteraciones excedido: N=%1.0f\n\n',N)
else
fprintf('\n\n x_%1.0f =\n\n',k);
fprintf('  %14.10f\n', u);
fprintf('\n')
end
```

Figura 8-13 •

Existen ciertos sistemas para los cuales el método de Richardson produce iteraciones convergentes a la única solución del sistema $A\vec{x} = \vec{b}$; para ello el siguiente teorema contiene condiciones suficientes, relativas a la matriz de coeficientes A. La demostración es sencilla y se deja de ejercicio al lector.

Teorema 8.6 *Sea $A = [a_{ij}]$ una matriz cuadrada de orden n tal que:*

1. $a_{ii} = 1$ *para todo* $i = 1, 2, \ldots, n$.
2. *A es diagonalmente dominante.*

Entonces las iteraciones del método de Richardson convergen a la única solución del sistema lineal $A\vec{x} = \vec{b}$ para cualquier condición inicial $\vec{x}_0$.

8.3.5 Método iterativo de Gauss-Seidel

En este apartado veremos el último método iterativo que trataremos en este libro para aproximar soluciones de sistemas lineales, el método de Gauss-Seidel. Introduciremos las ideas principales a través del siguiente ejemplo.

▶ **Ejemplo 8.36** Consideremos el sistema

$$\begin{aligned} 8x_1 - x_2 + x_3 &= -6 \\ 3x_1 + 9x_2 - 6x_4 &= -9 \\ x_1 - 6x_2 + 12x_3 &= 23 \\ -x_1 + x_2 - x_3 + 6x_4 &= 5 \end{aligned} \tag{8.41}$$

Al despejar sucesivamente las variables x_i de cada una de las ecuaciones, obtenemos

$$x_1 = \frac{1}{8}x_2 - \frac{1}{8}x_3 - \frac{3}{4} \tag{8.42}$$

$$x_2 = -\frac{1}{3}x_1 + \frac{2}{3}x_4 - 1 \tag{8.43}$$

$$x_3 = -\frac{1}{12}x_1 + \frac{1}{2}x_2 + \frac{23}{12} \tag{8.44}$$

$$x_4 = \frac{1}{6}x_1 - \frac{1}{6}x_2 + \frac{1}{6}x_3 + \frac{5}{6} \tag{8.45}$$

Ahora, si sustituimos el valor x_1 de (8.42) en (8.43), ésta se transforma, después de simplificar, en

$$x_2 = -\frac{1}{24}x_2 + \frac{1}{24}x_3 + \frac{2}{3}x_4 - \frac{3}{4} \tag{8.46}$$

De manera análoga, al sustituir los valores x_2 de (8.46) y x_1 de (8.42) en la ecuación (8.44), después de simplificar, se tiene

$$x_3 = -\frac{1}{32}x_2 + \frac{1}{32}x_3 + \frac{1}{3}x_4 + \frac{77}{48} \tag{8.47}$$

Finalmente, al sustituir los valores x_1 de (8.42), x_2 de (8.46) y x_3 de (8.47) en (8.45), esta última ecuación se convierte en

$$x_4 = \frac{13}{576}x_2 - \frac{13}{576}x_3 - \frac{1}{18}x_4 + \frac{317}{288} \tag{8.48}$$

Al conjuntar la información de (8.42), (8.46), (8.47) y (8.48), se obtiene el sistema

$$\begin{aligned} x_1 &= \frac{1}{8}x_2 - \frac{1}{8}x_3 - \frac{3}{4} \\ x_2 &= -\frac{1}{24}x_2 + \frac{1}{24}x_3 + \frac{2}{3}x_4 - \frac{3}{4} \\ x_3 &= -\frac{1}{32}x_2 + \frac{1}{32}x_3 + \frac{1}{3}x_4 + \frac{77}{48} \\ x_4 &= \frac{13}{576}x_2 - \frac{13}{576}x_3 - \frac{1}{18}x_4 + \frac{317}{288} \end{aligned}$$

equivalente al sistema lineal (8.41). Podemos escribir el último sistema en forma matricial:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{8} & -\frac{1}{8} & 0 \\ 0 & -\frac{1}{24} & \frac{1}{24} & \frac{2}{3} \\ 0 & -\frac{1}{32} & \frac{1}{32} & \frac{1}{3} \\ 0 & \frac{13}{576} & -\frac{13}{576} & -\frac{1}{18} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} -\frac{3}{4} \\ -\frac{3}{4} \\ \frac{77}{48} \\ \frac{317}{288} \end{bmatrix}.$$

Entonces, si $g(\vec{x}) = T\vec{x} + \vec{c}$, con

$$T = \begin{bmatrix} 0 & \frac{1}{8} & -\frac{1}{8} & 0 \\ 0 & -\frac{1}{24} & \frac{1}{24} & \frac{2}{3} \\ 0 & -\frac{1}{32} & \frac{1}{32} & \frac{1}{3} \\ 0 & \frac{13}{576} & -\frac{13}{576} & -\frac{1}{18} \end{bmatrix} \quad \text{y} \quad \vec{c} = \begin{bmatrix} -\frac{3}{4} \\ -\frac{3}{4} \\ \frac{77}{48} \\ \frac{317}{288} \end{bmatrix},$$

las soluciones del sistema lineal (8.41) coinciden con las soluciones de la transformación afín g.◀

El ejemplo precedente contiene la idea central del llamado método iterativo de Gauss-Seidel. Note que $\|T\|^{(\infty)} = \frac{3}{4} < 1$, por lo que las iteraciones de punto fijo $\vec{x}_k = g(\vec{x}_{k-1}) = T\vec{x}_{k-1} + \vec{c}$ convergen para cualquier valor inicial $\vec{x}_0$.

Consideremos ahora el caso general de un sistema lineal $A\vec{x} = \vec{b}$, $A = [a_{ij}]$, $\vec{b} = [b_1 \cdots b_n]^t$, $\vec{x} = [x_1 \cdots x_n]^t$; recapitulemos la parte algebraica del método expuesto en el ejemplo anterior:

(a) Se aplica la parte algebraica del método de Jacobi al sistema

$$A\vec{x} = \vec{b} \qquad \text{(I)}$$

es decir, se despeja la primera variable de la primera ecuación, la segunda variable de la tercera ecuación, etc., y se obtiene un sistema afín

$$\vec{x} = [t_{ij}]\vec{x} + \vec{c} \qquad \text{(II)}$$

el sistema afín del método de Jacobi.

(b) En la segunda ecuación del sistema (II) se sustituye el valor x_1 de la primera ecuación del sistema (II).

(c) En la tercera ecuación del sistema (II) se sustituye el valor x_1 de la primera ecuación y el valor x_2 de la ecuación que se obtuvo en el inciso anterior.

(d) En la cuarta ecuación del sistema (II) se sustituye el valor x_1 de la primera ecuación de (II), el valor x_2 de la ecuación que se obtuvo en el inciso *(b)*, y el valor x_3 que se obtuvo en el inciso *(c)*.

(e) El proceso se repite en forma recurrente hasta hacer lo propio con la variable x_n de la última ecuación de (II).

El procedimiento, aunque en teoría es viable, puede resultar sumamente laborioso. Sin embargo, lo que en realidad requerimos es el proceso iterativo para aproximar la solución del sistema lineal y éste se puede efectuar con una pequeña modificación al método de Jacobi; pues al tener valores concretos, sólo se deben ir sustituyendo en las demás relaciones del método de Jacobi inmediatamente que se van obteniendo.

▶ **Ejemplo 8.37** Vamos a encontrar, sin utilizar la matriz T, las primeras dos iteraciones de la relación de recurrencia $g(\vec{x}_k) = T\vec{x}_{k-1} + \vec{c}$ del ejemplo 8.36 para aproximar la solución del sistema lineal (8.41), con la condición inicial $\vec{x}_0 = (0,0,0,0)$. Para ello utilicemos el método de Jacobi contenido en las ecuaciones (8.42) a (8.45). Sustituyendo $\vec{x} = \vec{x}_0$ en (8.42), se tiene

$$x_1^{(1)} = \frac{1}{8}0 - \frac{1}{8}0 - \frac{3}{4} = -\frac{3}{4}.$$

Ahora, sustituyendo este valor por x_1; esto es $x_1 = x_1^{(1)}$ y $x_2 = 0$, $x_3 = 0$ y $x_4 = 0$, en (8.43),

$$x_2^{(1)} = -\frac{1}{3}(-\frac{3}{4}) + \frac{2}{3}0 - 1 = -\frac{3}{4}.$$

En el siguiente paso hacemos las sustituciones $x_1 = x_1^{(1)}$, $x_2 = x_2^{(1)}$, $x_3 = 0$ y $x_4 = 0$ en (8.44) para obtener

$$x_3^{(1)} = -\frac{1}{12}(-\frac{3}{4}) + \frac{1}{2}(-\frac{3}{4}) + \frac{23}{12} = \frac{77}{48}.$$

Haciendo $x_1 = x_1^{(1)}$, $x_2 = x_2^{(1)}$, $x_3 = x_3^{(1)}$ y $x_4 = 0$ en (8.45), produce

$$x_4^{(1)} = \frac{1}{6}(-\frac{3}{4}) - \frac{1}{6}(-\frac{3}{4}) + \frac{1}{6}(\frac{77}{48}) + \frac{5}{6} = \frac{317}{288}.$$

De manera análoga:

$$\begin{aligned}
x_1^{(2)} &= \frac{1}{8}x_2^{(1)} - \frac{1}{8}x_3^{(1)} - \frac{3}{4} \\
&= \frac{1}{8}(-\frac{3}{4}) - \frac{1}{8}(\frac{77}{48}) - \frac{3}{4} \\
&= -\frac{401}{384} \\
&\approx -1.044270833, \\
x_2^{(2)} &= -\frac{1}{3}x_1^{(2)} + \frac{2}{3}x_4^{(1)} - 1 \\
&= -\frac{1}{3}(-\frac{401}{384}) + \frac{2}{3}(\frac{317}{288}) - 1 \\
&= \frac{283}{3456} \\
&\approx 0.081886574 07 \\
x_3^{(2)} &= -\frac{1}{12}x_1^{(1)} + \frac{1}{2}x_2^{(1)} + \frac{23}{12} \\
&= -\frac{1}{12}(-\frac{401}{384}) + \frac{1}{2}(\frac{283}{3456}) + \frac{23}{12} \\
&= \frac{28265}{13824} \\
&\approx 2.044632523
\end{aligned}$$

$$\begin{aligned} x_4^{(2)} &= \frac{1}{6}x_1^{(2)} - \frac{1}{6}x_2^{(2)} + \frac{1}{6}x_3^{(2)} + \frac{5}{6} \\ &= \frac{1}{6}(-\frac{401}{384}) - \frac{1}{6}(\frac{283}{3456}) + \frac{1}{6}(\frac{28265}{13824}) + \frac{5}{6} \\ &= \frac{81817}{82944} \\ &\approx .9864125193 \blacktriangleleft \end{aligned}$$

Entonces, para el caso general si $\vec{x}_k = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$, el ciclo iterativo que produce el método de Gauss-Seidel está dado por la relación de recurrencia

$$x_i^{(k)} = \frac{1}{a_{ii}}\left[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k-1)}\right].$$

Resumimos lo precedente en el siguiente algoritmo.

Método de Gauss-Seidel con margen de tolerancia entre las diferencias de iteraciones sucesivas

Algoritmo 8.9

Entrada: *La matriz de coeficientes,* $A = [a_{ij}] \in \mathfrak{M}_n(\mathbb{R})$, *del sistema* $A\vec{x} = \vec{b}$, *donde* $a_{ii} \neq 0$ *para todo* i*; el término independiente* $\vec{b} = [b_1 \ \cdots \ b_n]^t$*; la aproximación inicial* $\vec{x}_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$*; la tolerancia* δ *para las diferencias entre iteraciones sucesivas y el máximo número de iteraciones* N.

Salida: *La primera iteración de punto fijo,* $\vec{x}_k = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$, *que satisface el criterio*

$$\|\vec{x}_k - \vec{x}_{k-1}\| < \delta$$

o el despliegue del texto "Máximo número de iteraciones excedido".

Paso 1. *Sea* $k = 1$.

Paso 2. *Mientras que* $k \leq N$ *siga los pasos 3 a 5.*

Paso 3. *Sean*

$$x_i^{(k)} = \frac{1}{a_{ii}}\left[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k-1)}\right] \tag{8.49}$$

y

$$\vec{x}_k = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$$

Paso 4. *Si* $\|\vec{x}_k - \vec{x}_{k-1}\|_\infty \geq \delta$, *sea* $k = k + 1$ *y se repiten los pasos 2 a 5.*

Paso 5. *Si* $\|\vec{x}_k - \vec{x}_{k-1}\| < \delta$, *se despliega la salida* $\vec{x}_k$.

Paso 6. *Si* $k = N$ *y* $\|\vec{x}_k - \vec{x}_{k-1}\| >= \delta$, *deplegar el mensaje* "Máximo número de iteraciones excedido".

Programa en MATLAB para el método de Gauss-Seidel

La figura 8-14 contiene el código en MATLAB del algoritmo 8.9. Las instrucciones de su uso están al inicio del guión; es recomendable que el lector pruebe una a una las instrucciones con un caso concreto para que pueda comprender por sí mismo el programa.

```
function Met_iter_Gauss_Seidel(A,b,x0,delta,N)
%***Met_iter_Gauss_Seidel(A,b,x0,delta,N)***
%calcula la primera iteración x_k,
%para aproximar la solución del sistema
%Ax=b, con iteraciones de punto fijo
%por el método de Gauss_Seidel,
%para la cual ||x_k-x_k-1||< delta
%con punto inicial x0 y un máximo
%de iteraciones N. Si la tolerancia
%no se cumple a las N iteraciones
%se despliega un mensaje.

clear x;
n=size(A,1);
k=1;
while k<=N
   x(1,k)=-(1/A(1,1))*A(1,2:n)*x0(2:n)+b(1)/A(1,1);
for i=2:n
   x(i,k)=-(1/A(i,i))*(A(i,1:i-1)*x(1:i-1,k)+A(i,i+1:n)*x0(i+1:n,:))+b(i)/A(i,i);
end
if norm(x(:,k)-x0,inf)>=delta
   x0=x(:,k);
   k=k+1;
else
   u=x(:,k);
   break
end end if k>N

fprintf('\n\n Número máximo de iteraciones excedido: N=%1.0f\n\n',N)
else
fprintf('\n\n x_%1.0f =\n\n',k);
fprintf('  %14.10f\n', u);
fprintf('\n')
end
```

Figura 8-14 •

► **Ejemplo 8.38** Aproximar la solución del sistema (8.41) del ejemplo 8.36 con un margen de tolerancia $\delta = 0.00001$ y con la condición inicial $\vec{x}_0 = (0,0,0,0)$.◄

Solución

```
EDU» A=[8 -1 1 0;3 9 0 -6;1 -6 12 0;-1 1 -1 6]
A =
      8     -1      1      0
      3      9      0     -6
      1     -6     12      0
     -1      1     -1      6
EDU» b=[-6 -9 23 5]';x0=[0 0 0 0]';
EDU» Met_iter_Gauss_Seidel(A,b,x0,.00001,30)
 x_8 =
   -1.0000001359
    0.0000002889
    2.0000001558
    0.9999999552
EDU»
```
✔

Forma matricial del método de Gauss-Seidel

Vamos a buscar una matriz T de tal manera que podamos expresar la relación de recurrencia (8.49) del método de Gauss-Seidel en la forma

$$\vec{x}_k = T\vec{x}_{k-1} + \vec{c}.$$

Multipliquemos cada una de las ecuaciones de (8.49) por a_{ii} para obtener el sistema

$$\begin{array}{lllllllllllll}
a_{11}x_1^{(k)} & & & & & = & - & a_{12}x_2^{(k-1)} & - & a_{13}x_3^{(k-1)} & - \cdots - & a_{1n}x_n^{(k-1)} & + \; b_1 \\
a_{21}x_1^{(k)} & + & a_{22}x_2^{(k)} & & & = & & & - & a_{23}x_3^{(k-1)} & - \cdots - & a_{2n}x_n^{(k-1)} & + \; b_2 \\
\vdots & & \ddots & & & \vdots & & & & & & & \vdots \;\; \vdots \\
a_{n1}x_1^{(k)} & + & a_{n2}x_2^{(k)} & + \cdots + & a_{nn}x_n^{(k)} & = & & & & & & & b_n
\end{array}$$

que en forma matricial se escribe

$$Q\vec{x}_k = -U\vec{x}_{k-1} + \vec{b}$$

donde

$$Q = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots \\ & & & \ddots & & \\ a_{(n-1)1} & a_{(n-1)2} & \cdots & \vdots & a_{(n-1)(n-1)} & 0 \\ a_{n1} & a_{n2} & a_{n3} & \cdots & & a_{nn} \end{bmatrix} \tag{8.50}$$

y

$$U = \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & a_{(n-1)n} \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{8.51}$$

Es decir, Q es la parte triangular inferior de A, incluyendo la diagonal, con las demás entradas nulas y U es la parte triangular superior de A, excluyendo la diagonal, con las restantes componentes nulas. Ya que hemos supuesto que los elementos de la diagonal de A son todos distintos de cero, Q es una matriz invertible y, por tanto,

$$\vec{x}_k = -(Q^{-1}U)\vec{x}_{k-1} + Q^{-1}\vec{b};$$

esto es,

$$\vec{x}_k = T\vec{x}_{k-1} + \vec{c} \tag{8.52}$$

donde $T = -Q^{-1}U$ y $\vec{c} = Q^{-1}\vec{b}$. Ahora bien,

$$U = A - Q$$

y, en consecuencia,

$$\begin{aligned} Q^{-1}U &= Q^{-1}(A - Q) \\ &= Q^{-1}A - I_n. \end{aligned}$$

Luego,

$$-Q^{-1}U = I_n - Q^{-1}A.$$

Así, la matriz que corresponde al formato general (8.33) (cfr. página 877) para la transformación afín (8.34) en el método de Gauss-Seidel es

$$T_Q = I_n - Q^{-1}A$$

y la transformación afín cuyos puntos fijos coinciden con las soluciones del sistema para este método es

$$g_Q(\vec{x}) = (I_n - Q^{-1}A)\vec{x} + \vec{c}$$

donde $\vec{c} = Q^{-1}\vec{b}$. Entonces el método iterativo de Gauss-Seidel será convergente si

$$\left\|I_n - Q^{-1}A\right\| < 1 \tag{8.53}$$

para alguna norma matricial natural. Esta condición es difícil de aplicar en la práctica, debido a la dificultad que entraña calcular Q^{-1}. El siguiente teorema contiene condiciones suficientes para la convergencia del método iterativo de Gauss-Seidel en términos de una característica muy simple de identificar en la matriz de coeficientes.

Teorema 8.7 *Sean A una matriz cuadrada de orden n y $\vec{b} \in \mathbb{R}^n$. Si la matriz A es diagonalmente dominante, entonces el método iterativo de Gauss-Seidel converge a la única solución del sistema lineal $A\vec{x} = \vec{b}$ para cualquier vector inicial $\vec{x}_0 \in \mathbb{R}^n$.*

DEMOSTRACIÓN ■ Sea Q la matriz (8.50), entonces el método iterativo de Gauss-Seidel en forma matricial está dado por (8.52). Sea λ un valor propio de la matriz $I_n - Q^{-1}A$ y $\vec{x} = (x_1, x_2, \ldots x_n)$ un vector propio correspondiente tal que $\|x\|_\infty = 1$. Entonces

$$(I_n - Q^{-1}A)\vec{x} = \lambda\vec{x}$$

que equivale a

$$\vec{x} - Q^{-1}Ax = \lambda\vec{x};$$

o

$$(Q - A)\vec{x} = \lambda Q\vec{x}.$$

Claramente $Q - A = -U$, donde U es la matriz (8.51); entonces, para cada $i = 1, 2, \ldots, n$,

$$-\sum_{j=i+1}^{n} a_{ij}x_j = \lambda \sum_{j=1}^{i} a_{ij}x_j.$$

Al despejar el último término del miembro derecho de la precedente igualdad obtenemos

$$\lambda a_{ii}x_i = -\sum_{j=i+1}^{n} a_{ij}x_j - \lambda \sum_{j=1}^{i-1} a_{ij}x_j \tag{8.54}$$

para cada $i = 1, 2, \ldots, n$. Sea m un subíndice donde $\vec{x}$ alcanza su magnitud respecto a la norma cúbica; esto es, $|x_m| = \|\vec{x}\| = 1$. Entonces, de (8.54), con $i = m$,

$$\begin{aligned}
|\lambda||a_{mm}| &= \left| -\sum_{j=m+1}^{n} a_{mj}x_j - \lambda \sum_{j=1}^{m-1} a_{mij}x_j \right| \\
&\le \left| \sum_{j=m+1}^{n} a_{mj}x_j \right| + |\lambda| \left| \sum_{j=1}^{m-1} a_{mj}x_j \right| \\
&\le \sum_{j=m+1}^{n} |a_{mj}||x_j| + |\lambda| \sum_{j=1}^{m-1} |a_{mj}||x_j| \\
&\le \sum_{j=m+1}^{n} |a_{mj}|\,\|\vec{x}\|_\infty + |\lambda| \sum_{j=1}^{m-1} |a_{mj}|\,\|\vec{x}\|_\infty \\
&= \left(\sum_{j=m+1}^{n} |a_{mj}| + |\lambda| \sum_{j=1}^{m-1} |a_{mj}| \right) \|\vec{x}\|_\infty \\
&= \sum_{j=m+1}^{n} |a_{mj}| + |\lambda| \sum_{j=1}^{m-1} |a_{mj}|.
\end{aligned}$$

De donde

$$|\lambda| \le \frac{\sum_{j=m+1}^{n} |a_{mj}|}{|a_{mm}| - \sum_{j=1}^{m-1} |a_{mj}|}.$$

Puesto que A es diagonalmente dominante

$$\sum_{j=1}^{m-1} |a_{mj}| + \sum_{j=m+1}^{n} |a_{mj}| < |a_{mm}|$$

y, por tanto,

$$\frac{\sum_{j=m+1}^{n} |a_{mj}|}{|a_{mm}| - \sum_{j=1}^{m-1} |a_{mj}|} < 1.$$

En consecuencia

$$|\lambda| < 1$$

para todo valor propio de la matriz $I_n - Q^{-1}A$; luego

$$\rho(I_n - Q^{-1}A) < 1.$$

De la precedente desigualdad y el teorema 8.5 (cfr. pág. 879), el método iterativo de Gauss-Seidel (8.52) converge a la única solución del sistema lineal $A\vec{x}$ para cualquier vector inicial $\vec{x}_0$. ■

8.3.6 Series de Neumann y método iterativo para aproximar la inversa de una matriz

En este apartado estudiaremos un método numérico iterativo para calcular la inversa de una matriz cuadrada. Los resultados serán básicamente los mismos que se trataron en el segmento 6.3.4 del tema "*Inversa de la matriz de Leontief*" del capítulo 6; salvo que aquí será en un contexto general; pues en aquel apartado nos restringimos a trabajar con la norma natural matricial $\|\cdot\|_1$ definida por 8.20 del ejemplo 8.16 y aquí trataremos cualquier norma matricial natural. Recomendamos al lector repasar estos resultados si es que ya los estudió o consultar el apartado 6.3.4 para ver las demostraciones de lo que aquí estableceremos sin ellas; pues son literalmente las mismas cambiando la norma $\|\cdot\|_1$ por una norma matricial natural cualquiera.

En lo que resta de esta subsección $\|\cdot\|$ representa una norma matricial natural que proviene de cierta norma vectorial denotada por el mismo símbolo.

Definición 8.6 **(Convergencia de sucesiones de matrices)** *Sea $\{A_\nu\}$ una sucesión de matrices $A_1, A_2, \ldots, A_\nu, \ldots$ en el espacio $\mathfrak{M}_n(\mathbb{R})$. Se dice que la sucesión $\{A_\nu\}$ converge o tiene como límite a una matriz $L \in \mathfrak{M}_n(\mathbb{R})$, cuando ν tiende a infinito y se escribe*

$$\lim_{\nu\to\infty} A_\nu = L,$$

si

$$\lim_{\nu\to\infty} \|A_\nu - L\| = 0.$$

Esto significa que dado cualquier $\varepsilon > 0$ existe $\nu_0 \in \mathbb{N}$, tal que

$$\|A_\nu - L\| < \varepsilon \quad \forall \nu \geq \nu_0.$$

Proposición 8.5 *Sea $\{A_\nu\}$ una sucesión de matrices. Entonces:*

1. *$\{A_\nu\}$ converge a lo más a una matriz L.*
2. *La sucesión $\{A_\nu\}$ es convergente si y sólo si*

$$\lim_{\nu,\mu\to\infty} \|A_\nu - A_\mu\| = 0.$$

Proposición 8.6 *Sean $\{A_\nu\}$ y $\{B_\nu\}$ un par de sucesiones de matrices de orden n que convergen a sendas matrices L y H. Sean $\alpha \in \mathbb{R}$ y $M \in \mathfrak{M}_n(\mathbb{R})$. Supongamos además que, para cada ν, $A_\nu = [a_{ij}^{(\nu)}]$ y $L = [l_{ij}]$. Entonces:*

1. $\lim_{\nu\to\infty} (A_\nu + B_\nu) = L + H$.
2. $\lim_{\nu\to\infty} \alpha A_\nu = \alpha L$.
3. $\lim_{\nu\to\infty} M A_\nu = ML$.
4. $\lim_{\nu\to\infty} A_\nu = L \Leftrightarrow \lim_{\nu\to\infty} a_{ij}^{(\nu)} = l_{ij} \quad \forall i, j$.
5. *Si $\|A\| < 1$, entonces* $\lim_{\nu\to\infty} A^\nu = 0$.

Teorema 8.8 **(Series de Neumann)** *Sea $A \in \mathfrak{M}_n(\mathbb{R})$ tal que $\|A\| < 1$. La serie*[12] *de Neumann (o serie geométrica de la matriz A) se define como*

$$S_\nu = I_n + A + A^2 + \cdots A^\nu$$

donde I_n es la matriz identidad de orden n. Entonces:

1. *La serie de Neumann converge; es decir,* $\lim_{\nu \to} S_\nu$ *existe.*
2. *La matriz $I_n - A$ es invertible.*
3. $$\lim_{\nu \to \infty} S_\nu = (I_n - A)^{-1}$$

   es decir,

$$\lim_{\nu \to \infty} (I_n + A + A^2 + \cdots A^\nu) = (I_n - A)^{-1} \quad (8.55)$$

El teorema 8.8 provee un método para aproximar la inversa de una matriz numéricamente; pues si se desea estimar B^{-1}, se puede tomar $B = I_n - A$ con $A = I_n - B$ en (8.55).

Corolario 8.3 *Sea B una matriz cuadrada de orden n. Si $\|I_n - B\| < 1$, entonces la matriz B es invertible y además*

$$B^{-1} = \lim_{\nu \to \infty} (I_n + A + A^2 + \cdots A^\nu),$$

donde $A = I_n - B$ e I_n es la matriz identidad de orden n.

▶ **Ejemplo 8.39** Consideremos la matriz

$$B = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix}.$$

Entonces

$$\begin{aligned} A &= I_2 - B \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix} \\ A &= \begin{bmatrix} 0.2 & -0.1 \\ -0.2 & 0.1 \end{bmatrix}. \end{aligned}$$

Puesto que $\|A\|^{(\infty)} = 0.3 < 1$, entonces

$$B^{-1} = \lim_{\nu \to \infty} \sum_{k=0}^{\nu} A^k$$

donde $A^0 = I_2$. Una aproximación para B^{-1} es $I_2 + A + A^2 + A^3$. Ya que

[12]Con más precisión: la sucesión de sumas parciales de la serie de Neumann.

$$A^2 = \begin{bmatrix} 0.06 & -0.03 \\ -0.06 & 0.03 \end{bmatrix},$$

$$A^3 = \begin{bmatrix} 0.018 & -0.009 \\ -0.018 & 0.009 \end{bmatrix},$$

tenemos

$$\begin{aligned} B^{-1} &\approx \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.2 & -0.1 \\ -0.2 & 0.1 \end{bmatrix} + \begin{bmatrix} 0.06 & -0.03 \\ -0.06 & 0.03 \end{bmatrix} + \begin{bmatrix} 0.018 & -0.009 \\ -0.018 & 0.009 \end{bmatrix} \\ &= \begin{bmatrix} 1.278 & -0.139 \\ -0.278 & 1.139 \end{bmatrix}. \end{aligned}$$

Al hacer el producto de la aproximación de B^{-1} con B, se obtiene

$$\begin{bmatrix} 1.278 & -0.139 \\ -0.278 & 1.139 \end{bmatrix} \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix} = \begin{bmatrix} 0.9946 & 0.0027 \\ 0.0054 & 0.9973 \end{bmatrix}.$$

que es próximo a la identidad. La estimación es más precisa si truncamos $\sum_{k=0}^{\nu} A^k$ en un valor ν más grande.◀

Ahora veamos cómo es posible escribir la sucesión de sumas parciales de la serie de Neumann,

$$S_\nu = \sum_{k=0}^{\nu} A^k,$$

en una relación de recurrencia. Para ello multipliquemos ambos lados de la igualdad

$$S_{k-1} = I_n + A + A^2 + \cdots + A^{k-1}$$

por la matriz A para obtener $AS_{k-1} = A + A^2 + \cdots + A^k$; entonces

$$S_k = I_n + AS_{k-1}.$$

Ya tenemos un método iterativo para calcular la inversa de $I_n - A$ y, por tanto, la inversa de una matriz B. Necesitamos además un criterio para detener el ciclo iterativo en un punto donde se considere que la estimación ya tiene un grado de precisión adecuado. Si la norma de A es suficientemente pequeña, de acuerdo con cierto margen de tolerancia, entonces la matriz A^k es muy cercana a la matriz $\mathscr{O}$, de esta manera se fija la tolerancia δ para detener el ciclo cuando $\|A\|^k < \delta$ por primera vez.

Método de series de Neumann para aproximar la inversa de una matriz

Algoritmo 8.10

Entrada: *La matriz B para aproximar su inversa y el margen de tolerancia $\delta > 0$.*

Salida: *La aproximación S_k de B^{-1} que se obtiene cuando por primera vez $\left(\|A\|^{(\infty)}\right)^k < \delta$ o el mensaje* "Posiblemente la matriz no es invertible".

Paso 1. *Calcular* $A = I_n - B$, *donde* I_n *es la identidad del mismo orden de* B.

Paso 2. *Calcular* $\|A\|^{(\infty)}$.

Paso 3. *Si* $\|A\|^{(\infty)} \geq 1$, *desplegar el mensaje* "Posiblemente la matriz no es invertible".

Paso 5. *Si* $\|A\|^{(\infty)} < 1$, *seguir los pasos 6 a 7.*

Paso 6. *Hacer* $S_0 = I_n$. *Mientras que* $\left(\|A\|^{(\infty)}\right)^k \geq \delta$, *calcular, para* $k = 1, 2, \ldots$

$$S_k = I_n + AS_{k-1}$$

Paso 7. *Cuando* $\left(\|A\|^{(\infty)}\right)^k < \delta$ *por primera vez, desplegar* $B^{-1} \approx S_k$.

Programa en MATLAB para aproximar la inversa de una matriz con la serie de Neumann

La figura 8-15 contiene el código de la función `Inversa_Neumann` que implementa en MATLAB el algoritmo 8.10 para aproximar la inversa de una matriz. Se trabaja nuevamente, en este programa, con la norma matricial natural $\|\cdot\|^{(\infty)}$. La descripción de la función viene en los comentarios al inicio del código. Es recomendable que el estudiante pruebe los comandos de esta función directamene para que por sí mismo logre comprender su funcionamiento.

```
function Inversa_Neumann(B,delta)
%***Inversa_Neumann(B,delta)***
%Calcula la primera aproximación, S_k,
%con la serie de Neumann para
%la inversa de B que cumple
%(||A||^inf)^k < delta, por primera vez,
%si ||I-B||^inf < 1.
n=size(B,1);
I=eye(n);
A=I-B;
S0=I;
k=1;
if norm(A,inf)>=1
   u=0;
   else
      while norm(A^k,inf)>=delta
         S0=I+A*(S0);
         k=k+1;
      end
   u=S0;
end
if u==0
   fprintf('\n')
   fprintf('norm(I-B,inf) = %4.2f >=1:\n\n',norm(A,inf));
   fprintf('Posiblemente la matriz no es invertible')
   fprintf('\n\n')
else
    fprintf('\n')
    fprintf(' S_%2.0f =\n',k);
    fprintf('\n')
  for j=1:n
    fprintf('%12.8f',u(j,:))
    fprintf('\n')
  end
   fprintf('\n')
end
```

Figura 8-15 • Función `Inversa_Neumann`.

► **Ejemplo 8.40** Utilizar la función `Inversa_Neumann` para aproximar la inversa de la matriz

$$B=\begin{bmatrix} 1 & 0.1 & -0.1 & 0.1 & 0.1\\ 0.3 & 1.1 & 0 & 0.1 & 0\\ 0.2 & 0 & 0.9 & -0.1 & 0\\ 0 & 0.1 & -0.1 & 1.2 & 0\\ -0.1 & 0 & 0 & 0.1 & 1.1 \end{bmatrix}$$

con una tolerancia $\delta = 0.00001$.◄

Solución

```
EDU» B=[1 .1 -.1 .1 .1;.3 1.1 0 .1 0;.2 0 .9 -.1 0;
0 0.1 -.1 1.2 0;-.1 0 0 .1 1.1]
EDU» Inversa_neumann(B,.00001)
S_11 =
  0.99567609 -0.08509045  0.10399997 -0.05967100 -0.09051580
 -0.27193465  0.93933715 -0.03696663 -0.06075473  0.02472173
 -0.22078762  0.01030715  1.09851356  0.10741042  0.02007163
  0.00426285 -0.07741801  0.09462273  0.84734913 -0.00038725
  0.09012855 -0.00069735  0.00085240 -0.08245611  0.90089744
EDU»
```
✔

► **Ejemplo 8.41** Utilizar la función `Inversa_Neumann` para aproximar la inversa de la matriz

$$B=\begin{bmatrix} 2 & 1 & 1\\ -3 & 1 & 2\\ 3 & 3 & 4 \end{bmatrix}$$

con una tolerancia $\delta = 0.001$.◄

Solución

```
EDU» B=[2 1 1;-3 1 2;3 3 4];
EDU» Inversa_Neumann(B,.001);
norm(I-B,inf) = 9.00 >=1:

Posiblemente la matriz no es invertible
EDU»
```
✔

La matriz B del ejemplo anterior es invertible como puede comprobar fácilmente el lector. En general, una matriz B puede ser invertible aun si $\|I_n - B\| \geq 1$. Con frecuencia en algunos de estos casos, cuando $\|I_n - B\| \geq 1$, los métodos directos —específicamente el método de Gauss— son la única opción y, de hecho, resultan no sólo la única alternativa sino la mejor. Veamos un caso importante en la siguiente proposición.

Proposición 8.7 *Sea $B = [b_{ij}]$ una matriz cuadrada de orden n diagonalmente dominante. Entonces B es invertible.*

DEMOSTRACIÓN ■ Supongamos que B no es invertible, entonces el sistema lineal $B\vec{x} = \vec{0}_{\mathbb{R}^n}$ tiene soluciones no triviales. Sea $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ una de ellas y $|\alpha_i| = \text{máx}_{1 \le j \le n} |\alpha_j|$. Entonces

$$b_{ii}\alpha_i = -\sum_{\substack{j=1 \\ j \neq i}}^{n} b_{ij}\alpha_j$$

y, por tanto,

$$\begin{aligned} |b_{ii}\alpha_i| &= |-\sum_{\substack{j=1 \\ j \neq i}}^{n} b_{ij}\alpha_j| \\ &\le \sum_{\substack{j=1 \\ j \neq i}}^{n} |b_{ij}\alpha_j| \\ &\le \sum_{\substack{j=1 \\ j \neq i}}^{n} |b_{ij}||\alpha_i| \\ &= |\alpha_i| \sum_{\substack{j=1 \\ j \neq i}}^{n} |b_{ij}| \end{aligned}$$

de donde

$$|b_{ii}| \le \sum_{\substack{j=1 \\ j \neq i}}^{n} |b_{ij}|;$$

lo cual es una contradicción a la hipótesis de que B es diagonalmente dominante. Luego, B debe ser invertible. ■

Más aún, el lector puede probar que si se aplica el método de Gauss, algoritmo 8.1, entonces no se requieren hacer intercambios de filas si la matriz es diagonalmente dominante. Luego, se puede aplicar el algoritmo 8.2 (pág. 828) para hallar la inversa de B sin necesidad de hacer intercambio alguno de renglones. De hecho, existe estabilidad en cuanto al crecimiento en los errores de redondeo para este método.

8.4 Transformaciones de Householder

Existen ciertas transformaciones ortogonales que son una herramienta fundamental en álgebra lineal numérica. Éstas se utilizan, por ejemplo, para transformar una matriz en un formato adecuado —con ceros en ciertas posiciones clave— y aplicar determinado método numérico reduciendo cálculos y uso de espacio en memoria para registro de componentes; para encontrar factorizaciones QR; y en una amplia variedad de aplicaciones que no tienen que ver con métodos numéricos. Nos referimos a las transformaciones de Householder y esta sección se dedica a su estudio; en la siguiente, veremos la gran utilidad que tienen en el cálculo numérico de valores y vectores propios de matrices.

8.4.1 Definiciones y transformaciones básicas

Definición 8.7 *Sea* $\vec{u} \in \mathbb{R}^n$ *un vector dado que satisface* $\vec{u} \cdot \vec{u} = 1$*; esto es,* $\vec{u}^t\vec{u} = 1$*. Una matriz de la forma*

$$H = I_n - 2\vec{u}\vec{u}^t,$$

donde I_n *es la identidad de orden n, se llama* ***matriz de Householder****; mientras que a la aplicación lineal* $\vec{x} \mapsto H\vec{x}$*,* $\vec{x} \in \mathbb{R}^n$*, se le dice* ***transformación de Householder****.*

Las matrices de Householder son simétricas y ortogonales; recordemos que una matriz Q es ortogonal si $Q^{-1} = Q^t$, que equivale a que las columnas de Q sean vectores ortonormales de $\mathbb{R}^n$ (cfr. teorema 4.14, pág. 282). Hacemos patente este importante resultado en la siguiente proposición.

Proposición 8.8 *Sea* $H = I_n - 2\vec{u}\vec{u}^t$ *una matriz de Householder, entonces:*

1. *H es simétrica.*
2. *H es ortogonal.*
3. $H^{-1} = H^t = H$.

DEMOSTRACIÓN ■ 1.

$$\begin{aligned} H^t &= (I_n - 2\vec{u}\vec{u}^t)^t \\ &= I_n^t - 2(\vec{u}\vec{u}^t)^t \\ &= I_n - 2(\vec{u}^t)^t\vec{u}^t \\ &= I_n - 2\vec{u}\vec{u}^t \\ &= H. \end{aligned}$$

2.

$$\begin{aligned} HH^t &= HH \\ &= (I_n - 2\vec{u}\vec{u}^t)(I_n - 2\vec{u}\vec{u}^t) \\ &= I_n - 2\vec{u}\vec{u}^t - 2\vec{u}\vec{u}^t + 4(\vec{u}\vec{u}^t)(\vec{u}\vec{u}^t) \\ &= I_n - 4\vec{u}\vec{u}^t + 4\vec{u}((\vec{u}^t\vec{u})\vec{u}^t) \\ &= I_n - 4\vec{u}\vec{u}^t + 4\vec{u}(1 \cdot \vec{u}^t) \\ &= I_n - 4\vec{u}\vec{u}^t + 4\vec{u}\vec{u}^t \\ &= I_n \end{aligned}$$

Por tanto, $H^t = H^{-1}$.

3. De los incisos anteriores se desprende $H^{-1} = H^t = H$. ■

Una matriz de Householder, $H = I_n - 2\vec{u}\vec{u}^t$, está completamente determinada por el vector $\vec{u}$. En lugar de registrar las n^2 componentes de la matriz H, únicamente se requiere registrar el vector $\vec{u}$. Al aplicar una transformación de Householder a un vector $\vec{x}$ o a una matriz A, se obtienen resultados muy simples de almacenar. En efecto,

$$\begin{aligned} H\vec{x} &= (I_n - 2\vec{u}\vec{u}^t)\vec{x} \\ &= \vec{x} - 2(\vec{u}\vec{u}^t)\vec{x} \\ &= \vec{x} - 2(\vec{u}^t\vec{x})\vec{u} \\ &= \vec{x} - 2\gamma\vec{u} \end{aligned}$$

donde

$$\gamma = \vec{u}^t \vec{x}$$

(es fácil probar que $(\vec{u}\vec{u}^t)\vec{x} = (\vec{u}^t\vec{x})\vec{u}$, ¡pruébelo!). Y, por tanto, si $\vec{a}_1$, $\vec{a}_2, \ldots, \vec{a}_n$, son las columnas de la matriz A, entonces (cfr. (1.2) del ejemplo 1.14, pág. 11)

$$\begin{aligned} HA &= \begin{bmatrix} H\vec{a}_1 & H\vec{a}_2 & \cdots & H\vec{a}_n \end{bmatrix} \\ &= \begin{bmatrix} \vec{a}_1 - 2\gamma_1\vec{u} & \vec{a}_2 - 2\gamma_2\vec{u} & \cdots & \vec{a}_n - 2\gamma_n\vec{u} \end{bmatrix} \end{aligned}$$

donde

$$\gamma_i = \vec{u}^t \vec{a}_i$$

para $i = 1, 2, \ldots, n$.

Con las funciones de Householder podemos transformar un vector dado en uno que tenga ceros en todas las coordenadas, excepto en la primera. Sea $\vec{x} = (x_1, x_2, \ldots, x_n) \neq \vec{0}_{\mathbb{R}}^n$ un vector dado, queremos encontrar una transformación de Householder, $\vec{y} \mapsto H\vec{y}$, y un número real α tal que $H\vec{x} = (\alpha, \underbrace{0, \ldots, 0}_{n-1})$.

Para determinar la transformación de Householder, $H = I_n - 2\vec{u}\vec{u}^t$, basta determinar $\vec{u} = (u_1, u_2, \ldots, u_n)$. Dado que H es ortogonal, entonces[13]

$$\begin{aligned} \|H\vec{x}\|^2 &= (H\vec{x})^t H\vec{x} \\ &= (\vec{x}^t(H^tH))\vec{x} \\ &= \vec{x}^t(I_n\vec{x}) \\ &= \vec{x}^t\vec{x} \\ &= \|\vec{x}\|^2. \end{aligned}$$

Luego, se debe tener

$$\begin{aligned} |\alpha| &= \|(\alpha, 0, \ldots, 0)\| \\ &= \|H\vec{x}\| \\ &= \|\vec{x}\|. \end{aligned}$$

Ya que H es su propia inversa,

$$\begin{aligned} \vec{x} &= H^{-1}H\vec{x} \\ &= H^{-1}(\alpha\vec{e}_1) \\ &= \alpha H\vec{e}_1 \\ &= \alpha(\vec{e}_1 - 2(\vec{u}^t\vec{e}_1)\vec{u} \\ &= \alpha(\vec{e}_1 - 2u_1\vec{u}). \end{aligned}$$

Entonces

$$\begin{aligned} x_1 &= \alpha(1 - 2u_1^2) \\ x_2 &= -2\alpha u_1 u_2 \\ &\vdots \\ x_n &= -2\alpha u_1 u_n. \end{aligned}$$

[13] Aquí, hemos representado la norma canónica de un vector, $\|(x_1, \ldots, x_n)\|_2 = (\sum_{i=1}^n x_i^2)^{1/2}$ por el símbolo $\|\cdot\|$ sin el subíndice 2.

De donde

$$u_1 = \pm\left(\frac{\alpha - x_1}{2\alpha}\right)^{1/2}$$

$$u_i = -\frac{x_i}{2\alpha u_1}, \quad i = 2, 3, \ldots, n.$$

Para que u_1 esté definido se requiere que $\alpha - x_1$ y α tengan el mismo signo; para ello bastaría con poner $\alpha = \|\vec{x}\|$. Sin embargo, para fines de estabilidad numérica, es más conveniente definir $\alpha = -\text{sign}(x_1)\,\|\vec{x}\|$, donde la función sign representa la aplicación

$$\text{sign}(a) = \begin{cases} 1 & \text{si } a \geq 0 \\ -1 & \text{si } a < 0 \end{cases}$$

Claramente α, así definida, satisface que $\alpha - x_1$ y α tengan los mismos signos y $|\alpha| = \|\vec{x}\|$. Entonces tenemos dos casos:

- $\alpha > 0$ ($x_1 < 0$). Elijamos

$$u_1 = -\left(\frac{\alpha - x_1}{2\alpha}\right)^{1/2}$$

entonces, si $\beta = \alpha(\alpha - x_1)$,

$$\begin{aligned} -2\alpha u_1 &= 2\alpha\left(\frac{\alpha - x_1}{2\alpha}\right)^{1/2} \\ &= 2\left(\frac{\alpha^2(\alpha - x_1)}{2\alpha}\right)^{1/2} \\ &= \frac{2}{\sqrt{2}}(\alpha(\alpha - x_1))^{1/2} \\ &= \sqrt{2\beta}. \end{aligned}$$

Así,

$$\begin{aligned} \vec{u} &= (u_1, u_2, \ldots, u_n) \\ &= (u_1, -\frac{x_2}{2\alpha u_1}, \ldots, -\frac{x_2}{2\alpha u_1}) \\ &= -\frac{1}{2\alpha u_1}(-2\alpha u_1^2, x_2, \ldots, x_n) \\ &= \frac{1}{\sqrt{2\beta}}(x_1 - \alpha, x_2, \ldots, x_n). \end{aligned}$$

Por tanto, si

$$\begin{aligned} \vec{v} &= (x_1 - \alpha, x_2, \ldots, x_n), \\ \|\vec{v}\| &= (\alpha^2 - 2\alpha x_1 + x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2} \\ &= (\alpha^2 - 2\alpha x_1 + \alpha^2)^{1/2} \\ &= (2\alpha(\alpha - x_1))^{1/2} \\ &= \sqrt{2\beta}. \end{aligned}$$

luego,

$$\vec{u} = \frac{1}{\sqrt{2\beta}}\vec{v}$$
$$= \frac{1}{\|\vec{v}\|}\vec{v}$$

y una transformación de Householder que satisface $H\vec{x} = (\alpha, 0, \ldots, 0)$ es

$$H\vec{y} = (I_n - 2(\vec{u}\vec{u}^{\,t}))\vec{y}$$

con

$$\alpha = \|\vec{x}\| = -\text{sign}(x_1)\,\|\vec{x}\|\,,$$
$$\vec{u} = \frac{1}{\|\vec{v}\|}\vec{v},$$
$$\vec{v} = (x_1 - \alpha, x_2, \ldots, x_n).$$

La matriz de Householder es entonces

$$H = I_n - 2\vec{u}\vec{u}^t$$
$$= I_n - \frac{1}{\beta}\vec{v}\vec{v}^t$$

con

$$\alpha = \|\vec{x}\|\,,$$
$$\vec{u} = \frac{1}{\|\vec{v}\|}\vec{v},$$
$$\vec{v} = (x_1 - \alpha, x_2, \ldots, x_n);$$

esto es, la matriz de Householder definida por las relaciones:

$$\left.\begin{array}{lcl} H & = & I_n - \frac{1}{\beta}\vec{v}\vec{v}^t \\ \alpha & = & -\text{sign}(x_1)\,\|\vec{x}\| \\ \beta & = & \alpha(\alpha - x_1) \\ \vec{v} & = & (x_1 - \alpha, x_2, \ldots, x_n) \end{array}\right\} \quad (8.56)$$

- Ahora supongamos $\alpha < 0$, es decir, $x_1 \geq 0$. Para este caso elijamos

$$u_1 = \left(\frac{\alpha - x_1}{2\alpha}\right)^{1/2}.$$

Entonces

$$-2\alpha u_1 = -2\alpha_1\left(\frac{\alpha - x_1}{2\alpha}\right)^{1/2}$$
$$= 2\left(\frac{\alpha^2(\alpha - x_1)}{2\alpha}\right)^{1/2}$$
$$= \frac{2}{\sqrt{2}}(\alpha(\alpha - x_1))^{1/2}$$

y, si nuevamente $\beta = \alpha(\alpha - x_1)$, se tiene otra vez

$$\vec{u} = \frac{1}{\sqrt{2\beta}}(x_1 - \alpha, x_2, \ldots, x_n).$$

Luego, la aplicación de Householder $\vec{y} \mapsto H\vec{y}$ que transforma al vector $\vec{x}$ en el vector $(\alpha, 0, \ldots, 0)$ está determinada también por las relaciones (8.56).

▶ **Ejemplo 8.42** Encontrar una transformación de Householder que transforme el vector $\vec{x} = (2, 6, 3)$ en un vector con las dos últimas coordenadas nulas.◀

Solución En este caso $\alpha = -\text{sign}(x_1)\|\vec{x}\| = -7$, $\vec{v} = (x_1 - \alpha, x_2, x_3) = (2+7, 6, 3) = (9, 6, 3)$, $\beta = 7(7+2) = 63$. Entonces

$$\begin{aligned} H &= I_3 - \frac{1}{\beta}\vec{v}\vec{v}^t \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{63}\begin{bmatrix} 9 \\ 6 \\ 3 \end{bmatrix}\begin{bmatrix} 9 & 6 & 3 \end{bmatrix} \\ &= \frac{1}{63}\begin{bmatrix} -18 & -54 & -27 \\ -54 & 27 & -18 \\ -27 & -18 & 54 \end{bmatrix} \end{aligned}$$

Comprobación:

$$\begin{aligned} \frac{1}{63}\begin{bmatrix} -18 & -54 & -27 \\ -54 & 27 & -18 \\ -27 & -18 & 54 \end{bmatrix}\begin{bmatrix} 2 \\ 6 \\ 3 \end{bmatrix} &= \frac{1}{63}\begin{bmatrix} -441 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -7 \\ 0 \\ 0 \end{bmatrix}. \quad ✔ \end{aligned}$$

En general, sea $\vec{x} = (x_1, x_2, \ldots, x_k, x_{k+1}, \ldots, x_n) \in \mathbb{R}^n$ un vector dado y supongamos que se desea encontrar una matriz de Householder H_k, tal que el vector $H_k\vec{x}$ tenga sus $n-k$ últimas coordenadas iguales a cero. Para ello, denotemos por I_{k-1} e I_{n-k+1} las matrices identidad de órdenes $k-1$ y $n-k+1$, respectivamente. Construyamos, por el método que se explicó antes, la matriz de Householder H_1 tal que

$$H_1\vec{x}_2 = (\alpha_k, \underbrace{0, \ldots, 0}_{n-k})$$

donde $\alpha_k = -\text{sign}(x_k)\|\vec{x}_2\|$, $\vec{x}_2 = (x_k, x_{k+1}, \ldots, x_n)$, $\vec{v}_k = (x_k - \alpha_k, x_{k+1}, \ldots, x_n)$, $\beta_k = \alpha_k(\alpha_k - x_k)$ y, por tanto,

$$H_1 = I_{n-k+1} - \frac{1}{\beta_k}\vec{v}_k\vec{v}_k^t.$$

Entonces, si $\vec{x}_1 = (x_1, \ldots, x_{k-1})$ y

$$H_k = \begin{bmatrix} I_{k-1} & \mathscr{O}_{(k-1),(n-k+1)} \\ \mathscr{O}_{(n-k+1),(k-1)} & H_1 \end{bmatrix},$$

donde $\mathscr{O}_{(k-1),(n-k+1)}$ y $\mathscr{O}_{(n-k+1),(k-1)}$ son las matrices cero de órdenes $(k-1)\times(n-k+1)$ y $(n-k+1)\times(k-1)$, respectivamente, se tiene

$$\begin{aligned} H_k\vec{x} &= \begin{bmatrix} I_{k-1} & \mathscr{O}_{(k-1),(n-k+1)} \\ \mathscr{O}_{(n-k+1),(k-1)} & H_1 \end{bmatrix} \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \end{bmatrix} \\ &= \begin{bmatrix} I_{k-1}\vec{x}_1 \\ H_1\vec{x}_2 \end{bmatrix} \\ &= (x_1,\ldots,x_{k-1},\alpha_k,0,0,\ldots,0). \end{aligned}$$

Además H_k es, efectivamente, una matriz de Householder; pues si $\vec{v} = \begin{bmatrix} \vec{0}_{\mathbb{R}^{n-k}} \\ \vec{v}_k \end{bmatrix}$ y $\vec{u} = (1/\|\vec{v}\|)\vec{v}$,

$$H_k = I_n - 2\vec{u}\vec{u}^t.$$

Note que $H_k\vec{x}$ es un vector que tiene las primeras $k-1$ coordenadas idénticas a las primeras $k-1$ coordenadas del vector $\vec{x}$ y las últimas $n-k$ componentes son nulas.

► **Ejemplo 8.43** Hallar una matriz de Householder tal que si $\vec{w} = (5,2,6,3)$, entonces la transformación $\vec{x} \mapsto H\vec{x}$ deja invariante la primera coordenada de $\vec{w}$ y las últimas dos coordenadas las transforma en ceros.◄

Solución Por el ejemplo 8.42 la matriz

$$H_1 = \frac{1}{63}\begin{bmatrix} -18 & -54 & -27 \\ -54 & 27 & -18 \\ -27 & -18 & 54 \end{bmatrix}$$

transforma el vector $\vec{w}_2 = (2,6,3)$ en el vector $H_1\vec{w} = \begin{bmatrix} -7 & 0 & 0 \end{bmatrix}$. Entonces, por lo ya explicado, la matriz

$$\begin{aligned} H &= \begin{bmatrix} 1 & \mathscr{O}_{1,3} \\ \mathscr{O}_{3,1} & H_1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{2}{7} & -\frac{6}{7} & -\frac{3}{7} \\ 0 & -\frac{6}{7} & \frac{3}{7} & -\frac{2}{7} \\ 0 & -\frac{3}{7} & -\frac{2}{7} & \frac{6}{7} \end{bmatrix} \end{aligned}$$

es Householder y

$$H\vec{w} = (5,-7,0,0).$$

Comprobación:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{2}{7} & -\frac{6}{7} & -\frac{3}{7} \\ 0 & -\frac{6}{7} & \frac{3}{7} & -\frac{2}{7} \\ 0 & -\frac{3}{7} & -\frac{2}{7} & \frac{6}{7} \end{bmatrix}\begin{bmatrix} 5 \\ 2 \\ 6 \\ 3 \end{bmatrix}\begin{bmatrix} 5 \\ -7 \\ 0 \\ 0 \end{bmatrix}. \checkmark$$

8.4.2 Factorización QR de Householder y sistemas lineales

Sea A una matriz real, cuadrada y de orden n. Sea $H_1 = I_n - \frac{1}{b_1}\vec{v}_1\vec{v}_1^t$, la matriz de Householder que transforma la primera columna de la matriz A en un múltiplo del vector $\vec{e}_1 = (1,\underbrace{0,\ldots,0}_{n-1})$. Entonces, si $\vec{a}_j$, $j = 1,2,\ldots,n$, son las columnas de la matriz A,

$$H_1A = [\ H_1\vec{a}_1 \quad H_1\vec{a}_2 \quad \cdots \quad H_1a_n\ ]$$
$$= \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1(n-1)}^{(1)} & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2(n-1)}^{(1)} & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & \cdots & a_{3(n-1)}^{(1)} & a_{3n}^{(1)} \\ 0 & a_{43}^{(1)} & a_{43}^{(1)} & \cdots & a_{4(n-1)}^{(1)} & a_{4n}^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & a_{n3}^{(1)} & \cdots & a_{n(n-1)}^{(1)} & a_{nn}^{(1)} \end{bmatrix}.$$

Aplicando ahora la matriz de Householder $H_2 = I_{n-1} - \frac{1}{b_2}\vec{v}_1\vec{v}_1^t$ para transformar la segunda columna de H_1A dejando invariante la primera componente y nulas las componentes tercera en adelante, obtendremos

$$H_2H_1A = \begin{bmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & \cdots & a_{1(n-1)}^{(2)} & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2(n-1)}^{(2)} & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3(n-1)}^{(2)} & a_{3n}^{(2)} \\ 0 & 0 & a_{43}^{(2)} & \cdots & a_{4(n-1)}^{(2)} & a_{4n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \cdots & a_{n(n-1)}^{2} & a_{nn}^{(2)} \end{bmatrix}.$$

Podemos continuar este proceso hasta producir una matriz triangular superior R:

$$H_{n-1}H_{n-2}\cdots H_2H_1A = \begin{bmatrix} a_{11}^{(n-1)} & a_{12}^{(n-1)} & a_{13}^{(n-1)} & \cdots & a_{1(n-1)}^{(n-1)} & a_{1n}^{(n-1)} \\ 0 & a_{22}^{(n-1)} & a_{23}^{(n-1)} & \cdots & a_{2(n-1)}^{(n-1)} & a_{2n}^{(n-1)} \\ 0 & 0 & a_{33}^{(n-1)} & \cdots & a_{3(n-1)}^{(n-1)} & a_{3n}^{(n-1)} \\ 0 & 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & a_{(n-1)(n-1)}^{(n-1)} & a_{(n-1)n}^{(n-1)} \\ 0 & 0 & 0 & \cdots & 0 & a_{nn}^{(n-1)} \end{bmatrix} = R.$$

Por tanto, ya que las matrices H_i son ortogonales y simétricas

$$\begin{aligned} A &= (H_1H_2\cdots H_{n-1})R. \\ &= QR \end{aligned}$$

donde $Q = H_1H_2 \cdots H_{n-1}$, por ser producto de matrices ortogonales,[14] es una matriz ortogonal y R es triangular superior; es decir, una factorización QR para la matriz A (cfr. teorema 4.13, pág. 279). Hemos probado así el siguiente teorema:

Teorema 8.9 **(Factorización QR de Householder)** *Sea A una matriz cuadrada de orden n con componentes reales. Entonces existe un par de matrices de orden n, Q y R, tales que*

$$A = QR$$

donde R es una matriz triangular superior y Q es producto de matrices de Householder y, por tanto, es una matriz ortogonal.

Al método de factorización que hemos descrito y resumido en el teorema precedente se le llama **factorización QR de Householder** para la matriz A.

De esta manera, un sistema

$$A\vec{x} = \vec{b}$$

se puede escribir como

$$QR\vec{x} = \vec{b}$$

y, por tanto,

$$\begin{aligned} R\vec{x} &= Q^{-1}\vec{b} \\ &= (H_{n-1} \cdots H_2H_1)\vec{b}. \end{aligned}$$

Después de calcular $Q^{-1}\vec{b} = (H_{n-1} \cdots H_2H_1)\vec{b}$, el sistema se puede resolver mediante sustitución regresiva.

► **Ejemplo 8.44** Resolver el sistema lineal

$$\begin{bmatrix} 1 & -1 & -3 \\ -2 & 4 & 3 \\ 2 & 3 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \\ 8 \end{bmatrix}$$

por medio de factorización QR de Householder.◄

Solución Calculemos H_1: en este caso

$\alpha = -\text{sign}(1)\,\|(1,-2,2)\| = -3$; $\beta = \alpha(\alpha - x_1) = 12$; $\vec{v} = (x_1 - \alpha, x_2, x_3) = (1+3,-2,2) = (4,-2,2)$;

y, por tanto,

[14] Si M_1 y M_2 son matrices ortogonales, entonces $(M_1M_2)(M_1M_2)^t = (M_1M_2)(M_2^tM_1^t) = M_1M_1^t = I$, por eso el producto de matrices ortogonales es una matriz ortogonal.

$$\begin{aligned} H_1 &= I_3 - \tfrac{1}{\beta}\vec{v}\vec{v}^t \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \tfrac{1}{12}\begin{bmatrix} 4 \\ -2 \\ 2 \end{bmatrix}\begin{bmatrix} 4 & -2 & 2 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}. \end{aligned}$$

Entonces

$$\begin{aligned} H_1A &= \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}\begin{bmatrix} 1 & -1 & -3 \\ -2 & 4 & 3 \\ 2 & 3 & 9 \end{bmatrix} \\ &= \begin{bmatrix} -3 & 1 & -3 \\ 0 & 3 & 3 \\ 0 & 4 & 9 \end{bmatrix}. \end{aligned}$$

Ahora trabajamos con la submatriz $B = \begin{bmatrix} 3 & 3 \\ 4 & 9 \end{bmatrix}$:

$\alpha = -\text{sign}(3)\,\|(3,4)\| = -5$; $\beta = \alpha(\alpha - x_1) = -5(-5-3) = 40$; $\vec{v} = (x_1 - \alpha, x_2) = (3+5,4) = (8,4)$;

y, por tanto,

$$\begin{aligned} H &= I_2 - \tfrac{1}{\beta}\vec{v}\vec{v} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \tfrac{1}{40}\begin{bmatrix} 8 \\ 4 \end{bmatrix}\begin{bmatrix} -1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix}. \end{aligned}$$

Luego,

$$H_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

y

$$\begin{aligned} R &= H_2H_1A \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{bmatrix}\begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}\begin{bmatrix} 1 & -1 & -3 \\ -2 & 4 & 3 \\ 2 & 3 & 9 \end{bmatrix} \\ &= \begin{bmatrix} -3 & 1 & -3 \\ 0 & -5 & -9 \\ 0 & 0 & 3 \end{bmatrix}. \end{aligned}$$

y

$$
\begin{aligned}
Q &= H_1H_2\\
&= \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3}\\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3}\\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3}\end{bmatrix}\begin{bmatrix} 1 & 0 & 0\\ 0 & -\frac{3}{5} & -\frac{4}{5}\\ 0 & -\frac{4}{5} & \frac{3}{5}\end{bmatrix}\\
&= \begin{bmatrix} -1/3 & 2/15 & -14/15\\ 2/3 & -2/3 & -1/3\\ -2/3 & -11/15 & 2/15\end{bmatrix}
\end{aligned}
$$

Así,

$$A = QR \Rightarrow \vec{b} = QR\vec{x} = H_2H_1\vec{x}$$

y, por tanto,

$$
\begin{aligned}
R\vec{x} &= Q^{-1}\vec{b}\\
&= H_2H_1\vec{b}\\
&= \begin{bmatrix} 1 & 0 & 0\\ 0 & -\frac{3}{5} & -\frac{4}{5}\\ 0 & -\frac{4}{5} & \frac{3}{5}\end{bmatrix}\begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3}\\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3}\\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3}\end{bmatrix}\begin{bmatrix} -1\\ -3\\ 8\end{bmatrix}\\
&= \begin{bmatrix} -7\\ -4\\ 3\end{bmatrix}.
\end{aligned}
$$

Esto es,

$$\begin{bmatrix} -3 & 1 & -3\\ 0 & -5 & -9\\ 0 & 0 & 3\end{bmatrix}\begin{bmatrix} x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix} -7\\ -4\\ 3\end{bmatrix}.$$

Al hacer sustitución regresiva al sistema escalonado anterior obtenemos

$$
\begin{aligned}
x_3 &= 1,\\
x_2 &= \frac{-49x_3}{-5} = -1,\\
x_1 &= \frac{-7 - x_2 + 3x_3}{-3} = 1;
\end{aligned}
$$

es decir, $\vec{x} = (1, -1, 1)$. ✔

Programa en MATLAB para el método de factorización QR de Householder

La figura 8-18 contiene la función `fac_QR_householder` para calcular, en MATLAB, la factorización QR de matrices cuadradas por el método de Householder. Esta utiliza internamente un par de funciones más: la función `Householder1` y la función `matriz_hh`. La primera, figura 8-17, calcula la matriz de Householder que transforma un vector dado $\vec{x} = (x_1, \ldots, x_n)$ en el vector $(\alpha, 0, \ldots, 0)$, donde $\alpha = -\text{sign}(x_1)\|\vec{x}\|$. La segunda función, figura 8-16*(a)*, obtiene, a partir de una matriz H de orden r

```
function L=matriz_HH(H,n)
r=length(H);
u=eye(n-r,n);
M1=[zeros(r,n-r) H];
L=[u;M1];
```

(a)

```
function z=signo(a)
if a>=0
   z=1;
else
   z=-1;
end
```

(b)

Figura 8-16 • Función `matriz_hh`.

```
function z=householder1(b)
%***householder1(b)****
%Calcula la matriz de Householder H
%tal que Hb=(alpha,0,...,0)
%donde alpha=-sign(b1)*||b||.
n=length(b);
I=eye(n);
x1=b(1);
alpha=-signo(b(1))*norm(b);
beta=alpha*(alpha-x1);
if beta==0
   z=I;
else
   v=b-alpha*I(:,1);
   z=I-(1/beta)*v*v';
end
```

Figura 8-17 • Función `Householder1`.

```
function [Q, R]=fac_QR_householder(A)
%***[Q, R]=fac_QR_householder(A)***
%Calcula los factores Q y R de la
%factorización A=QR por el
%método de factorización QR de
%Householder.
  n=length(A);
  B=A(:,1);
  H1=householder1(B);
  R1=H1*A;
  Q1=H1;
for j=1:n-2
   B=R1(:,j+1);
   B(1:j)=[];
   H2=householder1(B);
   H=matriz_HH(H2,n);
   R1=H*R1;
   Q1=Q1*H;
end
Q=Q1;
R=R1;
```

Figura 8-18 • Función `fac_QR_householder`.

y un número entero $n \geq r$, una matriz de la forma $\begin{bmatrix} I_{n-r} & \mathcal{O}_{n-r,r} \\ \mathcal{O}_{n-r,r} & H \end{bmatrix}$. Finalmente, el procedimiento `Householder1` utiliza la función `signo` de la figura 8-16*(b)*.

▶ **Ejemplo 8.45** Utilizar la función `fac_QR_householder` para encontrar la factorización QR del método de Householder de la matriz de coeficientes

$$A = \begin{bmatrix} 1 & -1 & -3 \\ -2 & 4 & 3 \\ 2 & 3 & 9 \end{bmatrix}$$

del sistema lineal del ejemplo 8.44◀

Solución

```
EDU» format rat
EDU»  A=[1 -1 -3;-2 4 3;2 3 9];
EDU»  [Q R]=fac_QR_householder(A)
Q =
    -1/3          2/15        -14/15
     2/3         -2/3          -1/3
    -2/3        -11/15          2/15

R =
     -3            1            -3
      0           -5            -9
      0            0             3

EDU»
```

✔

8.4.3 Reducción de Householder-Hessenberg

En este apartado estudiaremos un método para reducir una matriz a una forma similar con ciertas entradas nulas por debajo de la diagonal, la forma superior de Hessenberg, que utilizaremos más adelante en el cálculo numérico de valores propios.

Definición 8.8 *Sea $A = [a_{ij}]$ una matriz cuadrada real de orden n.*

1. *A está en forma superior de Hessenberg si $a_{ij} = 0$ para todo $i \geq j+2$.*
2. *A está en forma inferior de Hessenberg si $a_{ij} = 0$ para todo $j \geq i+2$.*
3. *A es una matriz tridiagonal si es una matriz en forma superior e inferior de Hessenberg.*

Es decir, una matriz A está en forma superior de Hessenberg si los elementos por debajo de la subdiagonal son todos nulos; está en forma inferior de Hessenberg si los elementos por encima de la supradiagonal son nulos; y es tridiagonal si los elementos debajo y por encima de la subdiagonal y la supradiagonal, respectivamente, son ceros.

► **Ejemplo 8.46** Sean

$$A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ -4 & 1 & 2 & 5 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & 3 & 2 \end{bmatrix}, B = \begin{bmatrix} 2 & 3 & 0 & 0 \\ -1 & 1 & -2 & 0 \\ 2 & 2 & -1 & 1 \\ 5 & -7 & 3 & 4 \end{bmatrix} \text{ y } C = \begin{bmatrix} 2 & 2 & 0 & 0 \\ 1 & 3 & -1 & 0 \\ 0 & 6 & 5 & 1 \\ 0 & 0 & 3 & -1 \end{bmatrix}.$$

Entonces A está en forma superior de Hessenberg, B está en forma inferior de Hessenberg y C es una matriz tridiagonal.◄

Veamos ahora cómo es posible, mediante transformaciones de Householder, encontrar una matriz ortogonal H y una matriz Z que está en forma superior de Hessenberg tal que $H^tAH = H^{-1}AH = Z$, para cualquier matriz cuadrada A. Para ello, utilizaremos la siguiente notación, si B es cualquier matriz cuadrada, $\vec{B}^{(j)}$ representa la columna j de la matriz B. Sea H_1 la matriz de Householder que transforma al vector $\vec{A}^{(1)} = (a_{11}, a_{21}, \ldots, a_{n1})$ en el vector $(a_{11}, \alpha_1, 0, \ldots, 0)$; por lo que analizamos en el apartado 8.4.1, pág. 902,

$$H_1 = \begin{bmatrix} 1 & \mathscr{O}_{1,n-1} \\ \mathscr{O}_{n-1,1} & \widehat{H}_1 \end{bmatrix}$$

donde $\widehat{H}_1$ es la matriz de Householder de orden $n-1$ que transforma al vector $(a_{21}, a_{31}, \ldots, a_{n1})$ en el vector $(\alpha_1, 0, \ldots 0)$; entonces

$$\begin{aligned} H_1A &= \begin{bmatrix} H_1\vec{A}^{(1)} & H_1\vec{A}^{(2)} & \cdots & H_1\vec{A}^{(n)} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & * & \cdots & * & * \\ \alpha_1 & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * & * \end{bmatrix} \end{aligned}$$

(donde los asteriscos representan ciertos números no necesariamente iguales entre sí) y, por tanto,

$$H_1AH_1 = \begin{bmatrix} a_{11} & * & \cdots & * & * \\ \alpha_1 & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * & * \end{bmatrix} \begin{bmatrix} 1 & \mathscr{O}_{1,n-1} \\ \mathscr{O}_{n-1,1} & \widehat{H}_1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,n-1}^{(1)} & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2,n-1}^{(1)} & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & \cdots & a_{3,n-1}^{(1)} & a_{3,n}^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{n,n-1}^{(1)} & a_{nn}^{(1)} \end{bmatrix}.$$

Sean $Z_1 = H_1AH_1$ y H_2 la matriz de Householder que transforma al vector $(a_{12}^{(1)}, a_{22}^{(1)}, a_{32}^{(1)}, \ldots, a_{n2}^{(1)})$ en el vector $(a_{12}^{(1)}, a_{22}^{(1)}, \alpha_2, 0, \ldots, 0)$, entonces, por lo visto en 8.4.1, pág. 902,

$$H_2 = \begin{bmatrix} I_2 & \mathscr{O}_{2,n-2} \\ \mathscr{O}_{n-2,2} & \widehat{H}_2 \end{bmatrix}$$

donde $\widehat{H}_2$ es una matriz de Householder de orden $n-2$ que transforma al vector $(a_{32}^{(1)}, a_{42}^{(1)}, \ldots, a_{n2}^{(1)})$ en el vector $(\underbrace{\alpha_2, 0, \ldots, 0}_{n-2})$; entonces

$$H_2Z = \begin{bmatrix} H_2\vec{Z}_1^{(1)} & H_2\vec{Z}_1^{(2)} & \cdots & H_2\vec{Z}_1^{(n)} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & * & * & \cdots & * \\ a_{21}^{(1)} & a_{22}^{(1)} & * & * & \cdots & * \\ 0 & \alpha_2 & * & * & \cdots & * \\ 0 & 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & * & \cdots & * \end{bmatrix}$$

donde, como antes, los asteriscos representan ciertos números no necesariamente iguales entre sí. Por tanto,

$$H_2Z_1H_2 = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & * & * & \cdots & * \\ a_{21}^{(1)} & a_{22}^{(1)} & * & * & \cdots & * \\ 0 & \alpha_2 & * & * & \cdots & * \\ 0 & 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & * & \cdots & * \end{bmatrix} \begin{bmatrix} I_2 & \mathscr{O}_{2,n-2} \\ \mathscr{O}_{n-2,2} & \widehat{H}_2 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & a_{14}^{(2)} & \cdots & a_{1n}^{(2)} \\ a_{21}^{(2)} & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & a_{32}^{(2)} & a_{33}^{(2)} & a_{34}^{(2)} & \cdots & a_{3n}^{(2)} \\ 0 & 0 & a_{43}^{(2)} & a_{44}^{(2)} & \cdots & a_{4n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & a_{n4}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}$$

Este proceso se puede continuar hasta obtener una matriz Z_{n-2} en forma superior de Hessenberg

$$Z_{n-2} = H_{n-2} \cdots H_2 H_1 A H_1 H_2 \cdots H_{n-2}$$

y, puesto que las matrices H_k son matrices de Householder, se sigue que el producto de ellas es ortogonal y si $H = H_1 H_2 \cdots H_{n-2}$, entonces

$$\begin{aligned} Z_{n-2} &= H^{-1} A H \\ &= H^t A H \end{aligned}$$

Hemos hecho plausible así la demostración del siguiente teorema y los detalles restantes de su demostración se dejan de ejercicio al lector.

Teorema 8.10 *Sea $A \in \mathfrak{M}_n(\mathbb{R})$ una matriz cualquiera. Entonces existe una matriz ortogonal H, producto de matrices de Householder, tal que la matriz*

$$Z = H^{-1} A H = H^t A H$$

está en forma superior de Hessenberg. En particular, si la matriz A es simétrica, entonces la matriz Z es una matriz tridiagonal simétrica.

Cuando se calcule por este método una matriz Z en forma superior de Hessenberg similar a una matriz A, diremos que Z es la reducción de la matriz A a la forma superior de Hessenberg.

► **Ejemplo 8.47** Sea

$$A = \begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & -1 & 1 & 1 \\ 2 & 1 & 1 & 3 \\ -1 & 2 & 1 & 1 \end{bmatrix}.$$

Entonces, si $\vec{u} = (2, 2, -1)$, utilizando en MATLAB la función `householder1` para hallar la matriz de Householder $\hat{H}_1$ que transforma el vector $\vec{u}$ en el vector $(\alpha_1, 0, 0)$, se obtiene

$$\hat{H}_1 = \begin{bmatrix} -0.6667 & -0.6667 & 0.3333 \\ -0.6667 & 0.7333 & 0.1333 \\ 0.3333 & 0.1333 & 0.9333 \end{bmatrix};$$

luego,

$$H_1 = \begin{bmatrix} 1.0000 & 0 & 0 & 0 \\ 0 & -0.6667 & -0.6667 & 0.3333 \\ 0 & -0.6667 & 0.7333 & 0.1333 \\ 0 & 0.3333 & 0.1333 & 0.9333 \end{bmatrix}$$

y, por tanto,

$$Z_1 = H_1 A H_1 = \begin{bmatrix} 1.0000 & -0.3333 & 2.2667 & 0.8667 \\ -3.0000 & -0.5556 & -1.4889 & -2.0889 \\ 0.0000 & -0.6889 & -0.7422 & 2.1378 \\ 0.0000 & -1.4889 & 0.1378 & 2.2978 \end{bmatrix}.$$

Al evaluar la función `householder1` en el vector $(-0.6889, -1.4889)$ para transformarlo en el vector $(\alpha_2, 0)$, resulta la matriz de Householder

$$\hat{H}_2 = \begin{bmatrix} -0.4199 & -0.9076 \\ -0.9076 & 0.4199 \end{bmatrix},$$

así que

$$H_2 = \begin{bmatrix} 1.0000 & 0 & 0 & 0 \\ 0 & 1.0000 & 0 & 0 \\ 0 & 0 & -0.4199 & -0.9076 \\ 0 & 0 & -0.9076 & 0.4199 \end{bmatrix}$$

y entonces

$$Z = H_2 H_1 A H_1 H_2 = H_2 Z_1 H_2 = \begin{bmatrix} 1.0000 & -0.3333 & -1.7384 & -1.6932 \\ -3.0000 & -0.5556 & 2.5210 & 0.4741 \\ 0.0000 & 1.6405 & 2.6289 & -1.4220 \\ 0.0000 & 0.0000 & 0.5780 & -1.0734 \end{bmatrix}$$

es una matriz, en forma superior de Hessenberg, similar a la matriz A.◀

Programa en MATLAB para el método de reducción de Householder-Hessenberg

La figura 8-19 contiene el código de la función `reduccion_hhh` para implementar en MATLAB el algoritmo de reducción de Householder-Hessenberg. El programa utiliza internamente las funciones `householder1` y `matriz_hh` de las figuras 8-17 y 8-16*(a)*, respectivamente, que se utilizaron para la programación del método de factorización QR del apartado precedente. La entrada es la matriz A que se va a reducir a forma superior de Hessenberg y la salida son un par de matrices, H y Z, tales que $H^t AH = Z$, donde H es una matriz ortogonal y Z es una matriz que está en forma superior de Hessenberg.

```
function [Z, P]=reduccion_hhh(A)
%*** Reducción Householder-Hessenberg ***
%reduccion_hhh(A)
%calcula un par de matrices, Z y P.
%Z es una matriz en forma superior
%de Hessenberg y P es una matriz
%ortogonal tal que P'*A*P=Z.
n=length(A);
Z1=A;
H1=eye(n);
for j=1:n-2
   u=Z1(j+1:n,j);
   h1=householder1(u);
   H2=matriz_HH(h1,n);
   Z1=H2*Z1*H2;
   H=H1*H2;
   H1=H;
end
Z=Z1;
P=H;
```

Figura 8-19 • Código de la función `reduccion_hhh`.

► **Ejemplo 8.48** Utilizar la función `reduccion_hhh` para calcular en MATLAB un par de matrices, H y Z, tales que la matriz $H^t AH = Z$ esté en forma superior de Hessenberg y H sea una matriz ortogonal si

$$A = \begin{bmatrix} 1 & 3 & -2 & 1 & -1 \\ 2 & -1 & 1 & 0 & 2 \\ -2 & 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 \\ -4 & 3 & 1 & 2 & 2 \end{bmatrix}.$$◀

Solución

```
EDU» A=[1 3 -2 1 -1;2 -1 1 0 2
-2 1 2 1 1;1 1 -1 1 1; -4 3 1 2 2];
EDU» [Z H]=reduccion_hhh(A)
Z =
    1.0000   -3.0000   -1.5206   -0.8757    1.7091
   -5.0000   -0.2000   -3.2692   -2.1010   -0.2405
    0.0000   -2.6306    1.3572    1.3561    0.5323
    0.0000    0.0000    1.4286    1.9583   -1.4957
    0.0000    0.0000    0.0000   -0.1424    0.8844
H =
    1.0000         0         0         0         0
         0   -0.4000   -0.8819   -0.0456    0.2452
         0    0.4000   -0.4106    0.0282   -0.8189
         0   -0.2000    0.0912   -0.9595   -0.1764
         0    0.8000   -0.2129   -0.2767    0.4880
EDU»
```

✔

► **Ejemplo 8.49** Encontrar la reducción de Hessenberg para la matriz simétrica utilizando en MATLAB la función `reduccion_hhh`.

$$A = \begin{bmatrix} 2 & -2 & 3 & 2 & 6 \\ -2 & 2 & 1 & -2 & -2 \\ 3 & 1 & 2 & 1 & 0 \\ 2 & -2 & 1 & -8 & 3 \\ 6 & -2 & 0 & 3 & 2 \end{bmatrix}$$ ◄

Solución

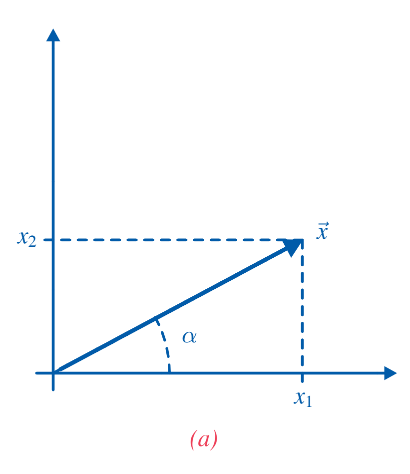
```
EDU» A=[2 -2 3 2 6;-2 2 1 -2 -2;3 1 2 1 0
2 -2 1 -8 3;6 -2 0 3 2];
EDU» reduccion_hhh(A)
ans =
    2.0000    7.2801    0.0000    0.0000    0.0000
    7.2801    3.8113    1.5054    0.0000    0.0000
    0.0000    1.5054    2.6653    1.6153    0.0000
    0.0000    0.0000    1.6153   -1.6910    4.1058
    0.0000    0.0000         0    4.1058   -6.7856
EDU»
```

✔

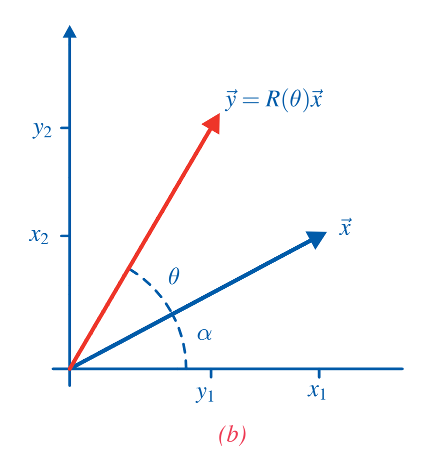
8.4.4 Rotaciones y reflexiones

Sea $\vec{x} = (x_1, x_2)$ un vector en el plano $\mathbb{R}^2$, como se ilustra en la figura 8-20*(a)*. Las coordenadas polares de $\vec{x}$, observando esta figura, vienen dadas por

$$\begin{aligned} x_1 &= r\cos\alpha, \\ x_2 &= r\,\mathrm{sen}\,\alpha, \end{aligned}$$



(a)



(b)

Figura 8-20 • Rotación de un vector.

con

$$r = \sqrt{x_1^2 + x_2^2},$$
$$\tan\alpha = \frac{x_2}{x_1}, \text{ si } x_1 \neq 0.$$

Y las obvias excepciones $\alpha = \pi/2$, si $x_1 = 0$ y $x_2 > 0$; $\alpha = \frac{3}{2}\pi$, si $x_1 = 0$ y $x_2 < 0$. Supongamos que el vector $\vec{x}$ se quiere rotar θ radianes preservando su longitud, como se ilustra en la figura 8-20*(b)*; obtendremos un nuevo vector $\vec{y} = (y_1, y_2)$ con

$$\begin{aligned}
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= \begin{bmatrix} r\cos(\theta+\alpha) \\ r\operatorname{sen}(\theta+\alpha) \end{bmatrix} \\
&= \begin{bmatrix} r(\cos\theta\cos\alpha - \operatorname{sen}\theta\operatorname{sen}\alpha) \\ r(\operatorname{sen}\theta\cos\alpha + \operatorname{sen}\alpha\cos\theta) \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} r\cos\alpha \\ r\operatorname{sen}\alpha \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.
\end{aligned}$$

De esta manera, si

$$R(\theta) = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

entonces la aplicación lineal $\vec{x} \mapsto R(\theta)\vec{x}$, transforma el vector $\vec{x}$ en un vector $\vec{y}$ que es la rotación de $\vec{x}$, preservando su magnitud, θ radianes. Dos casos particulares son cuando se rota el vector $\vec{x}$ y el resultado está sobre uno de los dos ejes. Si se desea que $\vec{y} = R\vec{x} = (r, 0)$, entonces $\theta + \alpha = 0$, i. e., $\theta = -\alpha$; por tanto, $\cos\theta = \cos\alpha = \frac{x_1}{r}$ y $\operatorname{sen}\theta = -\operatorname{sen}\alpha = -\frac{x_2}{r}$; así

$$R = \begin{bmatrix} \frac{x_1}{r} & \frac{x_2}{r} \\ -\frac{x_2}{r} & \frac{x_1}{r} \end{bmatrix} \tag{8.57}$$

pues

$$R\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \dfrac{x_1}{r} & \dfrac{x_2}{r} \\ -\dfrac{x_2}{r} & \dfrac{x_1}{r} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix};$$

De manera análoga el lector puede proceder para deducir que si $\vec{y} = R\vec{x} = (0, r)$, la matriz R está dada por

$$R = \begin{bmatrix} \frac{x_2}{r} & -\frac{x_1}{r} \\ \frac{x_1}{r} & \frac{x_2}{r} \end{bmatrix}.$$

La matriz $R(\theta)$ es ortogonal, ya que

$$(\cos\theta, \operatorname{sen}\theta)\cdot(-\operatorname{sen}\theta, \cos\theta) = 0$$

y

$$(\cos\theta, \operatorname{sen}\theta)\cdot(\cos\theta, \operatorname{sen}\theta) = 1,$$
$$(-\operatorname{sen}\theta, \cos\theta)\cdot(-\operatorname{sen}\theta, \cos\theta = 1.$$

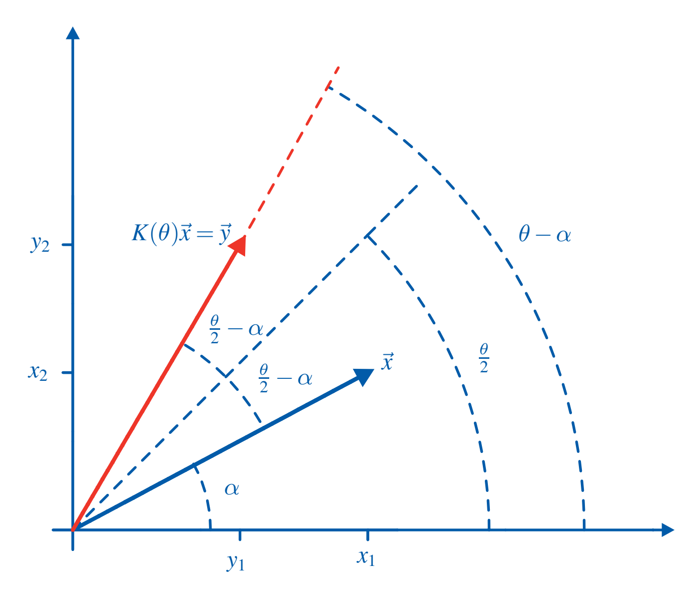


Figura 8-21 • Reflexión del vector $\vec{x}$ alrededor de la recta con pendiente $\theta/2$ y que pasa por el origen.

En la figura 8-21 el vector $\vec{y} = (y_1, y_2)$ se obtiene reflejando el vector $\vec{x} = (x_1, x_2)$ tomando como eje de reflexión la recta $x_2 = \tan(\theta/2)x_1$. Entonces, de esta figura se desprende

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} r\cos(\theta-\alpha) \\ r\operatorname{sen}(\theta-\alpha) \end{bmatrix} = \begin{bmatrix} r\cos\theta\cos\alpha + r\operatorname{sen}\theta\operatorname{sen}\alpha \\ r\operatorname{sen}\theta\cos\alpha - r\operatorname{sen}\alpha\cos\theta \end{bmatrix} = \begin{bmatrix} x_1\cos\theta + x_2\operatorname{sen}\theta \\ x_1\operatorname{sen}\theta - x_2\cos\theta \end{bmatrix}$$
$$= \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix};$$

luego, si

$$G(\theta) = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix},$$

la aplicación $\vec{x} \mapsto G(\theta)\vec{x}$ transforma el vector $\vec{x}$ en el vector $\vec{y}$ que es la reflexión del primero alrededor de la recta $x_2 = \tan(\theta/2)x_1$. La matriz G es evidentemente ortogonal y simétrica; más aún, es una matriz de Householder. En efecto: si $\vec{u} = (\operatorname{sen}(\theta/2), -\cos(\theta/2))$, entonces

$$\begin{aligned} I_2 - 2\vec{u}\vec{u}^t &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 2\begin{bmatrix} \operatorname{sen}(\theta/2) \\ -\cos(\theta/2) \end{bmatrix}\begin{bmatrix} \operatorname{sen}(\theta/2) & -\cos(\theta/2) \end{bmatrix} \\ &= \begin{bmatrix} 1 - 2\operatorname{sen}^2\frac{1}{2}\theta & 2\operatorname{sen}\frac{1}{2}\theta\cos\frac{1}{2}\theta \\ 2\operatorname{sen}\frac{1}{2}\theta\cos\frac{1}{2}\theta & 1 - 2\cos^2\frac{1}{2}\theta \end{bmatrix} \\ &= \begin{bmatrix} 1 - 2\dfrac{1-\cos\theta}{2} & \operatorname{sen}\theta \\ \operatorname{sen}\theta & 1 - 2\dfrac{1+\cos\theta}{2} \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix} \\ &= G(\theta). \end{aligned}$$

Un caso particular es hallar la matriz $G(\theta)$ para que el resultado de la reflexión del vector $\vec{x}$ esté sobre el eje x; esto es, $G\vec{x} = (r, 0)$. Entonces se debe tener $\theta - \alpha = 0$ y, por tanto, $\cos\theta = x_1/r$, $\operatorname{sen}\theta = x_2/r$; luego

$$G = \begin{bmatrix} \frac{x_1}{r} & \frac{x_2}{r} \\ \frac{x_2}{r} & -\frac{x_1}{r} \end{bmatrix} \tag{8.58}$$

▶ **Ejemplo 8.50** Si $\vec{x} = (-4, 3)$, hallar una matriz de rotación R que transforme este vector en un vector con la misma magnitud sobre el eje x; esto es, $R\vec{x} = (5, 0)$.◀

Solución En este caso $r = \sqrt{x_1^2 + x_2^2} = \sqrt{16 + 9} = 5$. Por (8.57)

$$R = \begin{bmatrix} \frac{x_1}{r} & \frac{x_2}{r} \\ -\frac{x_2}{r} & \frac{x_1}{r} \end{bmatrix} = \begin{bmatrix} -\frac{4}{5} & \frac{3}{5} \\ -\frac{3}{5} & -\frac{4}{5} \end{bmatrix}.$$

Comprobación:

$$\begin{bmatrix} -\frac{4}{5} & \frac{3}{5} \\ -\frac{3}{5} & -\frac{4}{5} \end{bmatrix}\begin{bmatrix} -4 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}. \quad ✔$$

▶ **Ejemplo 8.51** Hallar una matriz de reflexión G que al aplicarse al vector $\vec{x} = (3, 4)$ lo transforme en un vector con la misma magnitud sobre el eje x; esto es, $G\vec{x} = (5, 0)$.◀

Solución $r = \sqrt{3^2 + 4^2} = 5$. Por (8.58)

$$G = \begin{bmatrix} \frac{x_1}{r} & \frac{x_2}{r} \\ \frac{x_2}{r} & -\frac{x_1}{r} \end{bmatrix} = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix}.$$

Comprobación:

$$\begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}. \checkmark$$

Consideremos el caso general en el espacio $\mathbb{R}^n$; para ello vamos a definir generalizaciones de las matrices R y G que estudiamos en el plano $\mathbb{R}^2$. Sean $i, j \in \{1, 2, \ldots, n\}$ un par de índices con $i \neq j$; $R = [r_{ut}]$ la matriz que satisface

$$\left.\begin{aligned} r_{ii} &= r_{jj} = \cos\theta \\ r_{ji} &= \operatorname{sen}\theta \\ r_{ij} &= -\operatorname{sen}\theta \\ r_{ut} &= \begin{cases} 1, & \text{si } u = t \\ 0, & \text{si } u \neq t \end{cases} ; \text{ para } u, t \neq i, j. \end{aligned}\right\} \tag{8.59}$$

y $G = [k_{ut}]$ la matriz definida por las relaciones

$$\left.\begin{aligned} k_{ii} &= \cos\theta \\ k_{jj} &= -\cos\theta \\ k_{ji} &= k_{ij} = \operatorname{sen}\theta, \\ k_{ut} &= \begin{cases} 1, & \text{si } u = t \\ 0, & \text{si } u \neq t \end{cases} ; \text{ para } u, t \neq i, j. \end{aligned}\right\} \tag{8.60}$$

Diremos que R es una matriz de **rotación** y que G es una matriz de **reflexión** de Givens (o simplemente una matriz de Givens). Observe que ambas matrices son iguales a la matriz identidad de orden n excepto en cuatro componentes. Es fácil comprobar que las matrices de rotación son ortogonales y las de reflexión son ortogonales y simétricas; la demostración de estas afirmaciones se deja de ejercicio al lector.

▶ **Ejemplo 8.52** Si $n = 4$, $i = 2$ y $j = 3$, entonces, por (8.59)

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta & 0 \\ 0 & \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

y por (8.60)

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \operatorname{sen}\theta & 0 \\ 0 & \operatorname{sen}\theta & -\cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \blacktriangleleft$$

Veamos el efecto de estas matrices al multiplicarlas por un vector $\vec{x} = (x_1, x_2, x_3, x_4)$:

$$R\vec{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta & 0 \\ 0 & \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 \\ (\cos\theta)x_2 - (\operatorname{sen}\theta)x_3 \\ (\operatorname{sen}\theta)x_2 + (\cos\theta)x_3 \\ x_4 \end{bmatrix}$$

y

$$G\vec{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \operatorname{sen}\theta & 0 \\ 0 & \operatorname{sen}\theta & -\cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 \\ (\cos\theta)x_2 + (\operatorname{sen}\theta)x_3 \\ (\operatorname{sen}\theta)x_2 - (\cos\theta)x_3 \\ x_4 \end{bmatrix}.$$

En el caso general si $c = \cos\theta$, $s = \operatorname{sen}\theta$, $\vec{x} = (x_1, x_2, \ldots, x_n)$ e $i < j$, entonces es fácil probar, lo cual se deja de ejercicio al lector, que

$$R\vec{x} = (x_1, \ldots, x_{i-1}, cx_i - sx_j, x_{i+1}, \ldots, x_{j-1}, sx_i + cx_j, x_{j+1}, \ldots, x_n) \tag{8.61}$$

y

$$G\vec{x} = (x_1, \ldots, x_{i-1}, cx_i + sx_j, x_{i+1}, \ldots, x_{j-1}, sx_i - cx_j, x_{j+1}, \ldots, x_n) \tag{8.62}$$

Es decir, ambas transformaciones dejan invariantes todas las componentes x_t del vector $\vec{x}$ excepto las componentes i y j; las cuales son combinaciones lineales de x_i y x_j. En la práctica se buscan transformaciones que dejen invariantes algunas componentes y transformen en ceros otras. Si se requiere una rotación que transforme en cero la componente j y deje invariantes las demás, excepto x_i y x_j, necesitamos que en (8.61) $sx_i + cx_j = 0$, lo cual implica $x_j\cos\theta = -x_i\operatorname{sen}\theta$; esto es, cuando

$$\left.\begin{aligned} c &= \cos\theta = \frac{x_i}{\sqrt{x_i^2 + x_j^2}} \\ s &= \operatorname{sen}\theta = -\frac{x_j}{\sqrt{x_i^2 + x_j^2}} \end{aligned}\right\} \tag{8.63}$$

De manera análoga, si

$$\left.\begin{aligned} c &= \cos\theta = \frac{x_i}{\sqrt{x_i^2 + x_j^2}} \\ s &= \operatorname{sen}\theta = \frac{x_j}{\sqrt{x_i^2 + x_j^2}} \end{aligned}\right\} \tag{8.64}$$

entonces la j-ésima coordenada de $G\vec{x}$ en (8.62) será cero.

► **Ejemplo 8.53** Encontrar una matriz de rotación que transforme al vector $\vec{x} = (-1, 4, 3)$ en un vector cuya tercera componente sea cero y deje invariante la primera coordenada.◄

Solución Utilicemos las relaciones (8.63) con $i = 2$ y $j = 3$; entonces

$$c = \frac{4}{5} \text{ y } s = -\frac{3}{5}$$

luego

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{5} & \frac{3}{5} \\ 0 & -\frac{3}{5} & \frac{4}{5} \end{bmatrix}.$$

Comprobación:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{5} & \frac{3}{5} \\ 0 & -\frac{3}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \\ 0 \end{bmatrix}. \quad ✔$$

8.5 Aproximación de valores y vectores propios

El concepto de valores y vectores propios tiene una gran variedad de aplicaciones en las ciencias e ingeniería; sin embargo, ya que el cálculo directo de los valores propios de una matriz depende de hallar las raíces de un polinomio, rara vez en la práctica se pueden calcular en forma exacta. Esto sucede porque el problema de encontrar las raíces de un polinomio es sumamente complicado, pues únicamente existen fórmulas cerradas, con este fin, para polinomios de grado menor o igual a cuatro. Incluso hallar el polinomio característico de una matriz de orden mayor, puesto que requiere del cálculo de un determinante, puede resultar virtualmente imposible para cierto tipo de matrices. Aun teniendo a la mano el polinomio característico, los métodos numéricos existentes para aproximar sus raíces pueden presentar dificultades si el grado del polinomio es muy grande. Para resolver este tipo de problemas, existen métodos numéricos que se basan en técnicas propias del álgebra lineal que son muy precisos. En este segmento daremos una introducción de algunos de los métodos numéricos que se utilizan con gran frecuencia en la práctica para aproximar valores y vectores propios; que no pretende de ninguna manera ser exhaustiva, pero sí completa para los objetivos que se quieren alcanzar en este libro. Por desgracia no podremos examinar, debido a limitaciones en extensión, otros métodos que son sumamente importantes, pero los que veremos aquí serán suficientes no sólo para dar una amplia perspectiva del tema, sino para proveer al lector de una herramienta con la que pueda atacar con cierta eficacia este tipo de problemas y comprender, cuando así lo necesite, aquellas otras metodologías que no contiene este texto y que se encuentran en la literatura de álgebra lineal numérica y análisis numérico que recomendamos en la bibliografía.

8.5.1 *Método de la potencia*

Sea A una matriz real de orden n y supongamos que uno de sus valores propios es **dominante**; es decir, A tiene un valor propio λ_1 cuya magnitud es mayor que la magnitud de cualquiera de los demás valores propios de A. Así, podemos ordenar de acuerdo con su magnitud los valores propios de A; esto es,

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots |\lambda_n|.$$

Es evidente que si $A \in \mathfrak{M}_n(\mathbb{R})$ tiene un valor propio dominante éste es real y de multiplicidad algebraica uno. Sea $\vec{v}_i$ un vector propio correspondiente al valor propio λ_i, para cada $i = 1, \ldots, n$, y supongamos además que los vectores $\vec{v}_i$ son linealmente independientes; en consecuencia forman una base de $\mathbb{R}^n$. Sean $\vec{u}_0 \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$ y $\alpha_1, \alpha_2 \ldots, \alpha_n$, escalares tales que

$$\vec{u}_0 = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n.$$

Entonces

$$\begin{aligned} A\vec{u}_0 &= \alpha_1 A\vec{v}_1 + \alpha_2 A\vec{v}_2 + \cdots + \alpha_n A\vec{v}_n \\ &= \alpha_1\lambda_1\vec{v}_1 + \alpha_2\lambda_2\vec{v}_2 + \cdots + \alpha_n\lambda_n\vec{v}_n, \\ A^2\vec{u}_0 &= \alpha_1\lambda_1 A\vec{v}_1 + \alpha_2\lambda_2 A\vec{v}_2 + \cdots + \alpha_n\lambda_n A\vec{v}_n \\ &= \alpha_1\lambda_1^2\vec{v}_1 + \alpha_2\lambda_2^2\vec{v}_2 + \cdots + \alpha_n\lambda_n^2\vec{v}_n; \end{aligned}$$

y se puede probar por inducción que para todo $k = 1, 2, \ldots$

$$A^k\vec{u}_0 = \alpha_1\lambda_1^k\vec{v}_1 + \alpha_2\lambda_2^k\vec{v}_2 + \cdots + \alpha_n\lambda_n^k\vec{v}_n.$$

Factoricemos λ_1^k de la precedente expresión para obtener

$$A^k\vec{u}_0 = \lambda_1^k\left(\alpha_1\vec{v}_1 + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{v}_2 + \cdots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{v}_n\right) \tag{8.65}$$

de donde

$$\frac{1}{\lambda_1^k}A^k\vec{u}_0 = \left(\alpha_1\vec{v}_1 + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{v}_2 + \cdots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{v}_n\right)$$

y, puesto que $|\lambda_i/\lambda_1| < 1$ para todo $i = 2, \ldots, n$, se tiene

$$\lim_{k\to\infty}\frac{1}{\lambda_1^k}A^k\vec{u}_0 = \lim_{k\to\infty}\left(\alpha_1\vec{v}_1 + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{v}_2 + \cdots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{v}_n\right) = \alpha_1\vec{v}_1.$$

De esta manera, si $\alpha_1 \neq 0$, la sucesión $\left(\frac{1}{\lambda_1^k}A^k\vec{u}_0\right)$ converge a un vector propio correspondiente al valor propio λ_1. Luego, si k es grande, $A^k\vec{u}_0$ es aproximadamente un vector propio correspondiente a λ_1. Por otra parte, sea φ una función que a cada vector $\vec{u}_k = A^k\vec{u}_0$ le asigna una determinada componente no nula, $(\vec{u}_k)_j$, del vector $\vec{u}_k$ —por ejemplo, $\varphi(\vec{u}_k) = j$ puede ser el índice de la primera componente distinta de cero de $\vec{u}_k$— y consideremos los cocientes

$$\mu_k = \frac{(\vec{u}_{k+1})_j}{(\vec{u}_k)_j}$$

entonces, por (8.65),

$$\begin{aligned} \lim_{k\to\infty}\mu_k &= \lim_{k\to\infty}\frac{(\vec{u}_{k+1})_j}{(\vec{u}_k)_j} = \lim_{k\to\infty}\frac{\lambda_1^{k+1}\left(\alpha_1\vec{v}_1 + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k+1}\vec{v}_2 + \cdots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k+1}\vec{v}_n\right)_j}{\lambda_1^k\left(\alpha_1\vec{v}_1 + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{v}_2 + \cdots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{v}_n\right)_j} \\ &= \lim_{k\to\infty}\frac{\lambda_1^{k+1}\left(\alpha_1(\vec{v}_1)_j + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k+1}(\vec{v}_2)_j + \cdots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k+1}(\vec{v}_n)_j\right)}{\lambda_1^k\left(\alpha_1(\vec{v}_1)_j + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k(\vec{v}_2)_j + \cdots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k(\vec{v}_n)_j\right)} \\ &= \frac{\lambda_1\alpha_1(\vec{v}_1)_j}{\alpha_1(\vec{v}_1)_j} = \lambda_1 \end{aligned}$$

siempre que α_1 sea distinto de cero; pues $|\lambda_i/\lambda_1| < 1$ para todo i.

▶ **Ejemplo 8.54** Los valores propios de la matriz

$$A = \begin{bmatrix} -2 & -1 & -3 \\ -7 & 4 & -9 \\ -3 & 3 & -4 \end{bmatrix}$$

son 2, -3 y -1, como el lector puede fácilmente constatar, con valor propio dominante $\lambda_1 = -3$. Apliquemos el método precedente para aproximar a λ_1 utilizando el vector inicial $\vec{u}_0 = (1, -2, 2)$ y la primera componente de los vectores $A^k\vec{u}_0$ para calcular los cocientes μ_k. Entonces, ya que

$$A\vec{u}_0 = \begin{bmatrix} -6 \\ -33 \\ -17 \end{bmatrix}, A^2\vec{u}_0 = \begin{bmatrix} 96 \\ 63 \\ -13 \end{bmatrix}, A^3\vec{u}_0 = \begin{bmatrix} -216 \\ -303 \\ -47 \end{bmatrix}, A^4\vec{u}_0 = \begin{bmatrix} 876 \\ 723 \\ -73 \end{bmatrix},$$

$$A^5\vec{u}_0 = \begin{bmatrix} -2256 \\ -2583 \\ -167 \end{bmatrix}, A^6 u_0 = \begin{bmatrix} 7596 \\ 6963 \\ -313 \end{bmatrix} \text{ y } A^7\vec{u}_0 = \begin{bmatrix} -21216 \\ -22503 \\ -647 \end{bmatrix},$$

se tiene:

$$\mu_1 = \frac{(A^2\vec{u}_0)_1}{(A\vec{u}_0)_1} = \frac{96}{-6} = -16,$$

$$\mu_2 = \frac{(A^3\vec{u}_0)_1}{(A^2\vec{u}_0)_1} = \frac{-216}{96} = -2.25,$$

$$\mu_3 = \frac{(A^4\vec{u}_0)_1}{(A^3\vec{u}_0)_1} = \frac{876}{-216} \approx -4.05556,$$

$$\mu_4 = \frac{(A^5\vec{u}_0)_1}{(A^4\vec{u}_0)_1} = \frac{-2256}{876} \approx -2.57534,$$

$$\mu_5 = \frac{(A^6\vec{u}_0)_1}{(A^5\vec{u}_0)_1} = \frac{7596}{-2256} \approx -3.36702 \quad \text{y}$$

$$\mu_6 = \frac{(A^7\vec{u}_0)_1}{(A^6\vec{u}_0)_1} = \frac{-21216}{7596} \approx -2.79305$$

La tabla 8-4 contiene la continuación de este proceso realizado en una computadora hasta el cálculo de μ_{40}; en ella observamos que los valores μ_k tienden al valor propio $\lambda_1 = -3$. Por tanto, la aproximación para el vector propio correspondiente es, con cinco cifras significativas,

$$\frac{1}{(3)^{40}} A^{40}\vec{u}_0 \approx \begin{bmatrix} 10 \\ 10 \\ 0 \end{bmatrix}. \blacktriangleleft$$

1	-16.00000	11	-3.02909	21	-3.00050	31	-3.00001
2	-2.25000	12	-2.98081	22	-2.99967	32	-2.99999
3	-4.05556	13	-3.01288	23	-3.00022	33	-3.00000
4	-2.57534	14	-2.99145	24	-2.99985	34	-3.00000
5	-3.36702	15	-3.00572	25	-3.00010	35	-3.00000
6	-2.79305	16	-2.99620	26	-2.99993	36	-3.00000
7	-3.15215	17	-3.00254	27	-3.00004	37	-3.00000
8	-2.90472	18	-2.99831	28	-2.99997	38	-3.00000
9	-3.06604	19	-3.00113	29	-3.00002	39	-3.00000
10	-2.95706	20	-2.99925	30	-2.99999	40	-3.00000

Tabla 8-4 •

Un inconveniente para llevar a efecto este método es el hecho de que los valores de λ_1^k y $A^k\vec{u}_0$ pueden ser muy grandes o muy pequeños. Para remediar este problema se debe normalizar en forma conveniente en cada paso del proceso. Antes de establecer el algoritmo necesitamos de algunos preliminares.

Definición 8.9 *Sea, para cada $\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, $\text{máxc}(\vec{x})$ la primera componente de $\vec{x}$ con mayor valor absoluto; es decir, $x_i = \text{máxc}(\vec{x})$ si y sólo si $\|\vec{x}\|_\infty = |x_i|$ e i es el primer índice para el cual esto sucede.*

► **Ejemplo 8.55** Si $\vec{x} = (-3, 2, 1, 3)$, $\vec{y} = (1, 2, -2)$ y $\vec{z} = (1, -3, 4, 7)$, entonces

$$\text{máxc}(\vec{x}) = -3,\ \text{máxc}(\vec{y}) = 2 \text{ y } \text{máxc}(\vec{z}) = 7. ◄$$

La función máxc tiene dos propiedades básicas que serán útiles en la demostración del teorema 8.11, las hacemos patentes en el siguiente lema y su demostración se deja de ejercicio al lector.

Lema 8.1 *Sea la función de $\mathbb{R}^n$ en $\mathbb{R}$, $\vec{x} \mapsto \text{máxc}(\vec{x})$, entonces:*

1. $\text{máxc}(\alpha\vec{x}) = \alpha\,\text{máxc}(\vec{x})\ \forall\alpha \in \mathbb{R},\ \forall\vec{x} \in \mathbb{R}^n$.
2. *La función* máxc *es continua en todo punto de $\mathbb{R}^n$.*

Teorema 8.11 **(Método de la potencia)** *Sean A una matriz real de orden n que es diagonalizable con valores propios λ_i y vectores propios correspondientes $\vec{v}_i$ que forman una base de $\mathbb{R}^n$. Se supone que λ_1 es el valor propio dominante de la matriz A y que $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots |\lambda_n|$. Si $\vec{u}_0 = \sum_{i=1}^n \alpha_i\vec{v}_i \in \mathbb{R}^n$ es cualquier vector dado, con $\alpha_1 \neq 0$, y se definen por recurrencia las siguientes sucesiones*

$$\left.\begin{aligned} \vec{z}_k &= A\vec{u}_{k-1} \\ \mu_k &= \text{máxc}(\vec{z}_k) \\ \vec{u}_k &= \frac{1}{\mu_k}\vec{z}_k \end{aligned}\right\} \tag{8.66}$$

$k = 1, 2, \ldots$, entonces:

1. *Existe $c \in \mathbb{R} - \{0\}$ tal que*

$$\lim_{k\to\infty} \vec{u}_k = c\vec{v}_1$$

es decir, la sucesión $(\vec{u}_k)$ converge a un valor propio correspondiente al valor propio dominante λ_1

2. $$\lim_{k\to\infty} \mu_k = \lim_{k\to\infty} \text{máxc}(\vec{z}_k) = \lambda_1$$

DEMOSTRACIÓN ■ Vamos primero a probar, por inducción, que

$$\vec{u}_k = \frac{1}{\text{máxc}(A^k\vec{u}_0)} A^k\vec{u}_0 \tag{8.67}$$

$\forall k = 1, 2, \ldots$ En efecto, para $k = 1$,

$$\vec{u}_1 = \frac{1}{\text{máxc}(\vec{z}_1)}\vec{z}_1 = \frac{1}{\text{máxc}(A\vec{u}_0)}A\vec{u}_0$$

Sea $m > 1$ un entero y supongamos que la relación (8.67) es cierta para $k = m$, entonces

$$\begin{aligned}\vec{z}_{m+1} = A\vec{u}_m &= A\left(\frac{1}{\text{máxc}(A^m\vec{u}_0)}A^m\vec{u}_0\right)\\ &= \frac{1}{\text{máxc}(A^m\vec{u}_0)}AA^m\vec{u}_0 = \frac{1}{\text{máxc}(A^m\vec{u}_0)}A^{m+1}\vec{u}_0\end{aligned}$$

y, por el inciso 1 del lema 8.1,

$$\text{máxc}(\vec{z}_{m+1}) = \text{máxc}\left(\frac{1}{\text{máxc}(A^m\vec{u}_0)}A^{m+1}\vec{u}_0\right) = \frac{1}{\text{máxc}(A^m\vec{u}_0)}\text{máxc}\left(A^{m+1}\vec{u}_0\right);$$

luego

$$\vec{u}_{m+1} = \frac{1}{\text{máxc}(\vec{z}_{m+1})}\vec{z}_{m+1} = \frac{\text{máxc}(A^m\vec{u}_0)}{\text{máxc}\left(A^{m+1}\vec{u}_0\right)}\frac{1}{\text{máxc}(A^m\vec{u}_0)}A^{m+1}\vec{u}_0 = \frac{1}{\text{máxc}(A^{m+1}\vec{u}_0)}A^{m+1}\vec{u}_0\,.$$

Por tanto (8.67) es cierta para todo $k = 1, 2, \ldots$

1. Por (8.65), la relación (8.67), el inciso 1 del lema 8.1 y la continuidad de la función máxc (inciso 2, lema 8.1) se tiene

$$\begin{aligned}\lim_{k\to\infty}\vec{u}_k &= \lim_{k\to\infty}\frac{1}{\text{máxc}(A^k\vec{u}_0)}A^k\vec{u}_0\\
&= \lim_{k\to\infty}\frac{\lambda_1^k\left(\alpha_1\vec{v}_1+\alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{v}_2+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{v}_n\right)}{\text{máxc}\left(\lambda_1^k\left(\alpha_1\vec{v}_1+\alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{v}_2+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{v}_n\right)\right)}\\
&= \lim_{k\to\infty}\frac{\lambda_1^k\left(\alpha_1\vec{v}_1+\alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{v}_2+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{v}_n\right)}{\lambda_1^k\text{máxc}\left(\alpha_1\vec{v}_1+\alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{v}_2+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{v}_n\right)}\\
&= \lim_{k\to\infty}\frac{\left(\alpha_1\vec{v}_1+\alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{v}_2+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{v}_n\right)}{\text{máxc}\left(\alpha_1\vec{v}_1+\alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{v}_2+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{v}_n\right)}\\
&= \frac{\alpha_1\vec{v}_1}{\text{máxc}(\alpha_1\vec{v}_1)} = \frac{\alpha_1\vec{v}_1}{\alpha_1\text{máxc}(\vec{v}_1)} = \frac{1}{\text{máxc}(\vec{v}_1)}\vec{v}_1 = c\vec{v}_1\end{aligned}$$

con $c = 1/\text{máxc}(\vec{v}_1)$.

2. Puesto que las funciones $\vec{x} \mapsto A\vec{x}$ y $\vec{x} \mapsto \text{máxc}(\vec{x})$ son continuas (cfr. inciso 2 del lema 8.1) y el hecho de que $\text{máxc}(\alpha\vec{x}) = \alpha\,\text{máxc}(\vec{x})$, se desprende, como consecuencia del inciso anterior,

$$\begin{aligned}
\lim_{k\to\infty}\mu_k &= \lim_{k\to\infty}\text{máxc}(\vec{z}_k)\\
&= \lim_{k\to\infty}\text{máxc}(A\vec{u}_{k-1})\\
&= \text{máxc}\left(\lim_{k\to\infty}(A\vec{u}_{k-1})\right)\\
&= \text{máxc}\left(A\lim_{k\to\infty}(\vec{u}_{k-1})\right)\\
&= \text{máxc}(Ac\vec{v}_1) = c\,\text{máxc}(A\vec{v}_1) = c\,\text{máxc}(\lambda_1\vec{v}_1) = c\lambda_1\text{máxc}(\vec{v}_1)\\
&= \frac{1}{\text{máxc}(\vec{v}_1)}\lambda_1\text{maxc}(\vec{v}_1) = \lambda_1 \quad ■
\end{aligned}$$

```
 k  maxc(z_k)      u_1        u_2        u_3

 1  -33.00000    0.18182    1.00000    0.51515
 2   -2.90909    1.00000    0.65625   -0.13542
 3   -3.15625    0.71287    1.00000    0.15512
 4   -2.89109    1.00000    0.82534   -0.08333
 5   -2.94863    0.87340    1.00000    0.06465
 6   -2.94077    1.00000    0.91667   -0.04121
 7   -2.96248    0.94281    1.00000    0.02875
 8   -2.97187    1.00000    0.96182   -0.01904
 9   -2.98138    0.97429    1.00000    0.01287
10   -2.98720    1.00000    0.98282   -0.00858
11   -2.99146    0.98850    1.00000    0.00575
12   -2.99426    1.00000    0.99232   -0.00384
13   -2.99617    0.99487    1.00000    0.00256
14   -2.99744    1.00000    0.99658   -0.00171
15   -2.99829    0.99772    1.00000    0.00114
16   -2.99886    1.00000    0.99848   -0.00076
17   -2.99924    0.99899    1.00000    0.00051
18   -2.99949    1.00000    0.99932   -0.00034
19   -2.99966    0.99955    1.00000    0.00023
20   -2.99977    1.00000    0.99970   -0.00015
```

Tabla 8-5 • Veinte iteraciones del método de la potencia para aproximar el valor propio dominante de la matriz A del ejemplo 8.56. Las columnas u_1, u_2 y u_3 contienen las coordenadas de la aproximación $\vec{u}_k$ de un vector propio correspondiente.

▶ **Ejemplo 8.56** Apliquemos el método de la potencia del teorema 8.11 para aproximar el valor propio dominante de la matriz

$$A = \begin{bmatrix} -2 & -1 & -3 \\ -7 & 4 & -9 \\ -3 & 3 & -4 \end{bmatrix}$$

del ejemplo 8.54, con el mismo vector inicial $\vec{u}_0 = (1, -2, 2)$. Entonces

$$\begin{aligned}
z_1 &= A\vec{u}_0 = (-6, -33, -17),\\
\text{máxc}(\vec{z}_1) &= -33,\\
\vec{u}_1 &= \frac{1}{-33}\vec{z}_1 = (\frac{6}{33}, 1, \frac{17}{33}) \approx (.18182, 1.0, .51515);
\end{aligned}$$

$$\vec{z}_2 = A\vec{u}_1 = (-32/11, -21/11, 13/33) \approx (-2.90909, -1.90909, .39394),$$
$$\text{máxc}(\vec{z}_2) = -32/11 \approx -2.90909,$$
$$\vec{u}_2 = -\frac{11}{32}\vec{z}_1 = (1, 21/32, -13/96) \approx (1.0, .65625, -.13542);$$
$$\vec{z}_3 = A\vec{u}_2 = (-9/4, -101/32, -47/96) \approx (-2.25, -3.15625, -.48958),$$
$$\text{máxc}(\vec{z}_3) = -101/32 \approx -3.15625$$
$$\vec{u}_3 = -\frac{32}{101}(-9/4, -101/32, -47/96) \approx (.712871, 1.0, .155116)$$

La tabla 8-5 contiene las primeras 20 iteraciones de este proceso. En ella se puede observar que la sucesión $\text{maxc}(\vec{z}_k)$ tiende al valor propio dominante $\lambda_1 = -3$ y el vector $\vec{u}_k$ tiende al vector propio correspondiente $(1, 1, 0)$.◀

Programa en MATLAB para el método de la potencia

La figura 8-22 contiene el algoritmo en código MATLAB para el método de la potencia. En él se ha incluido un criterio para detener el programa cuando por primera vez $\|\vec{u}_k - \vec{u}_{k-1}\|_\infty$ esté dentro de una tolerancia especificada por el usuario. También se emite un mensaje para excluir divisiones entre cero en el caso de que la matriz tenga un valor propio nulo. Las entradas de este programa son la matriz A, el vector inicial $\vec{u}_0$, el margen de tolerancia para $\|\vec{u}_k - \vec{u}_{k-1}\|_\infty$ y un número máximo de iteraciones permisible;

```
function metodo_potencia(A,u,eps,m);
%***Método de la potencia***
%metodo_potencia(A,u,eps,m)
%aproxima el valor dominante de la matriz A
%con un vector de iteración inicial u
%y margen de tolerancia eps para
%||u_k-u_(k-1)||_inf
%La salida son las iteraciones mínimas
%para aproximar el valor propio dominante
%y la aproximación a un vector propio correspondiente
x=u;
d=1;
K=0;
while (d>eps)& (K<m)
   K=K+1;
   [z j]=max(abs(x));
   y=A*x;
   R=y(j);
   [w t]=max(abs(y));
   if y(t)==0
      fprintf('\nEl escalar 0 es un valor propio de la matriz,\n')
      fprintf('cambiar el vector inicial\n')
   break
   else
   d=norm(x-y/y(t),inf);
   x=y/y(t);
   fprintf(' %3d %10.5f\n',K,R)
   end
end
fprintf('\n Vector propio correspondiente:\n   x = ')
fprintf('\n %14.6f\n',x)
```

Figura 8-22 •

mientras que la salida consiste en el mínimo de iteraciones necesarias para aproximar el valor propio dominante y la aproximación para un vector propio correspondiente.

▶ **Ejemplo 8.57** Utilizar la función `metodo_potencia` en MATLAB para aproximar el valor propio dominante de la matriz A del ejemplo precedente comenzando con el mismo vector inicial $\vec{u}_0$, con un margen de tolerancia $\|\vec{u}_k - \vec{u}_{k-1}\|_\infty < .000001$ y un máximo de 100 iteraciones.◀

Solución

```
EDU» A=[-2 -1 -3;-7 4 -9;-3 3 -4];u=[1;-2;2];
EDU» metodo_potencia(A,u,.00001,100);u=[1;-2;2];
   1  -33.00000
   2   -1.90909
   3   -2.25000
   4   -2.38614
   5   -2.57534
   6   -2.69570
   7   -2.79305
   8   -2.85842
   9   -2.90472
  10   -2.93587
  11   -2.95706
  12   -2.97127
  13   -2.98081
  14   -2.98719
  15   -2.99145
  16   -2.99430
  17   -2.99620
  18   -2.99746
  19   -2.99831
  20   -2.99887
  21   -2.99925
  22   -2.99950
  23   -2.99967
  24   -2.99978
  25   -2.99985
  26   -2.99990
  27   -2.99993
  28   -2.99996
  29   -2.99997
  30   -2.99998

 Vector propio correspondiente:
   x =
       1.000000

       0.999995

      -0.000003
EDU»
```
✔

▶ **Ejemplo 8.58** Utilizar el programa `metodo_potencia` en MATLAB para aproximar el valor propio dominante de la matriz

$$B = \begin{bmatrix} 1 & -1 & 4 \\ -1 & -1 & -2 \\ -1 & 0 & -3 \end{bmatrix}$$

utilizando el vector inicial $\vec{u}_0 = (3, -1, 1)$.◀

Solución

```
EDU» B=[1 -1 4;-1 -1 -2;-1 0 -3];u=[3 -1 -1]';
EDU» metodo_potencia(B,u,.00001,100)
El escalar 0 es un valor propio de la matriz,
cambiar el vector inicial
Vector propio correspondiente:
   x =
       3.000000

      -1.000000

      -1.000000

EDU» metodo_potencia(B,[1 -1 1]',.00001,100)

   1    6.00000
   2   -1.33333
   3   -1.50000
   4   -1.66667
   5   -1.80000
   6   -1.88889
   7   -1.94118
   8   -1.96970
   9   -1.98462
  10   -1.99225
  11   -1.99611
  12   -1.99805
  13   -1.99902
  14   -1.99951
  15   -1.99976
  16   -1.99988
  17   -1.99994
  18   -1.99997
  19   -1.99998

 Vector propio correspondiente:
   x =
       1.000000

      -0.999992

      -0.999996
 EDU»
```

✔

8.5.2 Deflación

El método de la potencia, como se presentó en el apartado anterior, se puede aplicar para encontrar un solo valor propio de una matriz cuadrada A de orden n, el valor propio dominante λ_1. Supongamos, sin embargo, que existe una matriz invertible H tal que

$$HAH^{-1} = \left[\begin{array}{c|ccc} \lambda_1 & b_2 & \cdots & b_n \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right] \tag{8.68}$$

donde A_1 es una matriz de orden $n-1$. Puesto que A y HAH^{-1} son similares, tienen los mismos valores propios; y ya que

$$\det(HAH^{-1} - \lambda I_n) = (\lambda_1 - \lambda)\det(A_1 - \lambda I_{n-1})$$

se desprende que los restantes $n-1$ valores propios de la matriz A son los valores propios de la matriz A_1; luego, se puede aplicar el método de la potencia para hallar el valor propio dominante λ_2 de la matriz A_1. Ahora bien, si una matriz H satisface la igualdad (8.68), se debe cumplir

$$HAH^{-1}\vec{e}_1 = \lambda_1\vec{e}_1,$$

donde $\vec{e}_1 = (1,0,\ldots,0)$ es el vector unitario de la base canónica de $\mathbb{R}^n$, que equivale a

$$\begin{aligned} A(H^{-1}\vec{e}_1) &= H^{-1}(\lambda_1\vec{e}_1) \\ &= \lambda_1(H^{-1}\vec{e}_1) \end{aligned}$$

y, por tanto, el vector $H^{-1}\vec{e}_1$ es un vector propio correspondiente al valor propio λ_1. Sea $\vec{x}$ ese vector; esto es,

$$\vec{x} = H^{-1}\vec{e}_1,$$

entonces

$$H\vec{x} = \vec{e}_1\,. \tag{8.69}$$

Es decir, una condición necesaria para que una matriz no singular cumpla la igualdad (8.68), es que exista un vector propio $\vec{x}$ correspondiente al valor propio λ_1 que satisfaga (8.69); veamos ahora que esta condición es suficiente. En efecto:

$$\begin{aligned} H\vec{x} = \vec{e}_1 &\Rightarrow \vec{x} = H^{-1}\vec{e}_1 \\ &\Rightarrow A\vec{x} = AH^{-1}\vec{e}_1 \\ &\Rightarrow \lambda_1\vec{x} = AH^{-1}\vec{e}_1 \\ &\Rightarrow \lambda_1 H^{-1}\vec{e}_1 = AH^{-1}\vec{e}_1 \\ &\Rightarrow H^{-1}(\lambda_1\vec{e}_1) = AH^{-1}\vec{e}_1 \\ &\Rightarrow \lambda_1\vec{e}_1 = HAH^{-1}\vec{e}_1. \end{aligned}$$

Ahora supongamos que existe un vector propio $\vec{u}$ correpondiente al valor propio λ_1 y que H es una matriz no singular tal que $H\vec{u} = -\vec{e}_1$, entonces $H^{-1}(-\vec{u}) = \vec{e}_1$ y $\vec{x} = -\vec{u}$ es un vector propio correspondiente a λ_1 y, por tanto, H cumple también (8.68). Resta únicamente encontrar una matriz invertible H que satisfaga la relación (8.69). Para lograr este fin, sea $\vec{y}$ un vector propio correspondiente al valor propio λ_1, y sea el vector unitario (respecto a la norma euclidiana $\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}}$)

$$\vec{x} = \frac{1}{\|\vec{y}\|}\vec{y},$$

entonces $\vec{x}$ es un vector propio correspondiente a λ_1 y podemos utilizar la matriz de Householder H de las fórmulas 8.56 (pág. 905) para que la función de Householder $\vec{u} \mapsto H\vec{u}$ transforme el vector $\vec{x} = (x_1, x_2, \ldots, x_n)$ en el vector $(-\text{sign}(x_1)\,\|\vec{x}\|, 0, \ldots 0) = -\text{sign}(x_1)\vec{e}_1$; específicamente:

$$\left.\begin{aligned} H &= I_n - \frac{1}{\beta}\vec{v}\vec{v}^t \\ \alpha &= -\text{sign}\,(x_1) \\ \beta &= 1 + \text{sign}(x_1)x_1 \\ \vec{v} &= (x_1 + \text{sign}(x_1), x_2, \ldots, x_n) \end{aligned}\right\} \tag{8.70}$$

Este es el llamado **método de deflación** de Householder e ilustramos su aplicación a continuación.

► **Ejemplo 8.59** Utilizar el método de deflación de Householder para encontrar los valores propios de la matriz

$$A = \begin{bmatrix} -2 & -1 & -3 \\ -7 & 4 & -9 \\ -3 & 3 & -4 \end{bmatrix}$$

del ejemplo 8.56.◄

Solución En ese ejemplo encontramos un vector propio correspondiente al valor propio dominante $\lambda_1 = -3$, $\vec{y} = (1,1,0)$. Luego, aplicando las relaciones (8.70) al vector

$$\vec{x} = \frac{1}{||\vec{y}||}\vec{y} = (1/\sqrt{2}, 1/\sqrt{2}, 0)$$

obtenemos

$$\alpha = -1\,,\ \beta = 1 + \frac{1}{\sqrt{2}}\,,\ \vec{v} = \begin{bmatrix} \frac{1}{\sqrt{2}} + 1 \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \quad \text{y}$$

$$\begin{aligned} H &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{1+\frac{1}{\sqrt{2}}} \begin{bmatrix} \frac{1}{\sqrt{2}} + 1 \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} + 1 & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} & 0 \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

que es la matriz de Householder que satisface $H\vec{x} = (1,0,0)$. Entonces, ya que[15]

$$\begin{aligned} HAH &= \begin{bmatrix} -\frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} & 0 \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & -1 & -3 \\ -7 & 4 & -9 \\ -3 & 3 & -4 \end{bmatrix} \begin{bmatrix} -\frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} & 0 \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -3 & -6 & 6\sqrt{2} \\ 0 & 5 & -3\sqrt{2} \\ 0 & 3\sqrt{2} & -4 \end{bmatrix}, \end{aligned}$$

los restantes valores propios de la matriz A son los valores propios de la matriz

$$A_1 = \begin{bmatrix} 5 & -3\sqrt{2} \\ 3\sqrt{2} & -4 \end{bmatrix}$$

[15]Recuerde que toda matriz de Householder es su propia inversa.

los cuales se pueden calcular directamente hallando las raíces del polinomio característico

$$p_{A_1}(\lambda) = \lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1);$$

es decir, $\lambda_2 = 2$ y $\lambda_3 = -1$. ✔

▶ **Ejemplo 8.60** Encontrar los valores propios de la matriz

$$A = \begin{bmatrix} -89 & 172 & -81 & 82 \\ -67 & 131 & -62 & 63 \\ -14 & 28 & -12 & 14 \\ 15 & -30 & 15 & -14 \end{bmatrix}$$

utilizando el método de deflación de Householder.◀

Solución Al utilizar la función `metodo_potencia` del programa de la figura 8-22 para encontrar el valor propio dominante de la matriz A, MATLAB produce $\lambda_1 = 16.000$ con vector propio correspondiente $\vec{y}_1 = (1.0, 0.8, 0.2, -0.2)$. Así, $\vec{x} = (1/\|\vec{y}\|)\vec{y} = (0.7625, 0.6100, 0.1525, -0.1525)$. Al aplicar (8.70) se obtiene

$$\alpha = -1,\ \beta_1 = 1 + 0.7625 \text{ y } \vec{v}_1 = (0.7625 + 1, 0.6100, 0.1525, -0.1525)$$

y, por tanto,

$$H_1 = \begin{bmatrix} -0.7625 & -0.6100 & -0.1525 & 0.1525 \\ -0.6100 & 0.7889 & -0.0528 & 0.0528 \\ -0.1525 & -0.0528 & 0.9868 & 0.0132 \\ -0.1525 & 0.0528 & 0.0132 & 0.9868 \end{bmatrix}.$$

Luego,

$$H_1AH_1 = \begin{bmatrix} 16.0000 & -253.5277 & 95.2928 & -96.8178 \\ 0.0000 & -5.6628 & 1.6673 & -1.5412 \\ 0.0000 & -7.1175 & 3.6789 & -1.8974 \\ 0.0000 & 4.7714 & -0.7654 & 1.9839 \end{bmatrix}.$$

Así,

$$A_1 = \begin{bmatrix} -5.6628 & 1.6673 & -1.5412 \\ -7.1175 & 3.6789 & -1.8974 \\ 4.7714 & -0.7654 & 1.9839 \end{bmatrix}.$$

Aplicando nuevamente la función `metodo_potencia` a la matriz A, MATLAB produce $\lambda_2 = -3$ con un vector propio correspondiente $\vec{y}_2 = (1.0000, 0.8299, -0.8299)$. Al normalizar, $\vec{x}_2 = (0.6485, 0.5382, -0.5382)$, entonces al utilizar otra vez (8.70),

$$\alpha = -1,\ \beta_2 = 1 + 0.6485 \text{ y } \vec{v}_1 = (0.6485 + 1, 0.5382, -0.5382)$$

y, por ende,

$$H_2 = \begin{bmatrix} -0.6485 & -0.5382 & 0.5382 \\ -0.5382 & 0.8243 & 0.1757 \\ 0.5382 & 0.1757 & 0.8243 \end{bmatrix};$$

luego

$$H_2 A_1 H_2 = \begin{bmatrix} -3.0000 & -7.7411 & 7.3563 \\ 0.0000 & 2.6471 & -1.0324 \\ 0.0000 & 1.0324 & 0.3529 \end{bmatrix}$$

y en consecuencia,

$$A_2 = \begin{bmatrix} 2.6471 & -1.0324 \\ 1.0324 & 0.3529 \end{bmatrix}.$$

Finalmente, los valores propios de A_2 se pueden calcular fácilmente mediante el polinomio característico

$$p_{A_2}(\lambda) = (\lambda - 2)(\lambda - 1).$$

Así que los valores propios de la matriz A son: $\lambda_1 = 16$, $\lambda_2 = -3$, $\lambda_3 = 2$ y $\lambda_4 = 1$. ✔

```
function [L, X]=met_potenciaII(A,u,eps,m);
x=u;
d=1;
K=0;
while (d>eps)& (K<m)
   K=K+1;
   [z j]=max(abs(x));
   y=A*x;
   R=y(j);
   [w t]=max(abs(y));
   if y(t)==0
      fprintf('\nEl escalar 0 es un valor propio de la matriz,\n')
      fprintf('cambiar el vector inicial\n')
   break
   else
   d=norm(x-y/y(t),inf);
   x=y/y(t);
   %fprintf(' %3d %10.5f\n',K,R)
   end
end
L=R;
X=x;
```

Figura 8-23 •

Programa en MATLAB para el método de deflación de Householder

La figura 8-24 contiene el código en MATLAB para calcular los valores propios de una matriz empleando el método de deflación de Householder que se explicó antes. Este programa necesita de dos funciones para su ejecución; la primera, `met_potenciaII` (figura 8-23), esencialmente es la función `metodo_potencia` con pequeñas adecuaciones; mientras que la segunda es la función `householder1`

```
function Lambdas=deflacion_hh(A,eps,m)
***Método de deflación de Householder***
%function Lambdas=deflacion_hh(A,eps,m)
%Calcula los valores propios de la matriz A
%con el método de deflación de Householder
%con una tolerancia eps y máximmo número
%de iteraciones m
n=length(A);
u=ones(n,1);
A1=A;
for j=1:n-1
   [lambda Y]=met_potenciaII(A1,u,eps,m);
   L(j)=lambda;
   X=(1/norm(Y))*Y;
   H=householder1(X);
   B=H*A1*H;
   B(1,:)=[];
   B(:,1)=[];
   A1=B;
   u=ones(n-j,1);
end
Lambdas=[L A1];
```

Figura 8-24 • Método de deflación de Householder.

de la figura 8-17 que se construyó antes en el método de factorización de Householder (cfr. pág 912). Las entradas del programa `deflacion_hh` son la matriz A, la tolerancia `eps` y el máximo número de iteraciones permisible `m`; mientras que la salida es el conjunto de los valores característicos. Se ha elegido, por conveniencia, un vector inicial con todas sus componentes iguales a la unidad en todo el ciclo. Esto puede tener eventualmente inconvenientes para algún caso particular; de ocurrir así, se tendrá que modificar el programa manualmente. Queda de tarea al lector mejorar el programa `deflacion_hh` para que, de manera interactiva, el usuario pueda cambiar el vector de inicio si es necesario.

▶ **Ejemplo 8.61** Utilizar el programa `deflacion_hh` para encontrar los valores propios de la matriz

$$A=\begin{bmatrix} -38-\sqrt{3} & -38-2\sqrt{3} & 36 & -24 \\ 19+\sqrt{3} & 19+2\sqrt{3} & -18 & 12 \\ -31 & -31 & 30 & -22 \\ -5 & -5 & 5 & -3 \end{bmatrix}$$ ◀

Solución

```
EDU» A=[-38-sqrt(3) -38-2*sqrt(3)  36 -24
19+sqrt(3) 19+2*sqrt(3) -18 12 -31 -31 30 -22 -5 -5 5 -3];
EDU»  deflacion_hh(A,.000001,100)
ans =
    7.0001    2.0000    1.7320    -1.0000
```
✔

❍ **Nota 8.6** 1. Para que el método de la potencia se pueda aplicar a una matriz, debe cumplir las condiciones del teorema 8.11; es decir, ser diagonalizable y poseer un valor propio dominante que, por definición, es real. Éste es un inconveniente para poder utilizar el método; pues se necesita saber *a priori* que la matriz de entrada satisface esos requisitos y, por desgracia, en la práctica difícilmente se puede hacer sin calcular antes los valores y vectores propios.

2. La hipótesis del valor propio dominante se puede debilitar un poco de la manera siguiente: si una matriz A es diagonalizable y tiene un valor propio real λ_1 tal que los valores propios λ_i de A satisfacen

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_m| > |\lambda_{m+1}| \ldots > |\lambda_n|,$$

en el método de las potencia las sucesiones $(\text{máxc}(z_k))$ y $(\vec{u}_k)$ aún convergen, respectivamente, a λ_1 y a un vector propio correspondiente. La demostración de este hecho se deja de ejercicio al lector.
3. El método de deflación se basa en el método de la potencia, entonces no garantiza resultados correctos si éste no se puede aplicar a cada submatriz que produce el ciclo.
4. Se puede probar, con las hipótesis del teorema 8.11, que el método de la potencia tiene "velocidad" de convergencia $(|\lambda_2|/|\lambda_1|)^k$; esto es, la convergencia será más rápida —requiere menos iteraciones para una buena aproximación— si $|\lambda_2|/|\lambda_1|$ es mucho menor que 1, y la convergencia es más lenta —requiere de mucho más iteraciones para una buena aproximación— si $|\lambda_2|/|\lambda_1|$ es cercano a 1.

8.5.3 Iteración inversa

El método de deflación de Householder únicamente aproxima los valores propios de una matriz para la cual es aplicable dicho método, pero no los vectores propios correspondientes; para ello existe un algoritmo, llamado **iteración inversa,** que sirve para aproximar un vector propio correspondiente a un valor propio. Para ello supongamos que λ_i, $i = 1, 2, \ldots n$, son los valores propios de una matriz diagonalizable $A \in \mathfrak{M}_n(\mathbb{R})$, con vectores propios correspondientes $\vec{v}_i$, que $\lambda_1 \in \mathbb{R}$ es un valor propio de multiplicidad algebraica uno y que $\rho \in \mathbb{R}$ es una muy buena aproximación de λ_1 —en el sentido de que $0 < |\lambda_1 - \rho| \ll |\lambda_i - \rho|$ para $i = 2, 3, \ldots, n$ (el símbolo $\ll$ significa "mucho menor que")—. Es fácil demostrar que la matriz $A - \rho I_n$ tiene por valores propios a $\lambda_1 - \rho, \lambda_2 - \rho, \ldots, \lambda_n - \rho$, con sendos vectores propios correspondientes $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$; y ya que los valores $\lambda_i - \rho$ son distintos a cero, la matriz $A - \rho I_n$ es invertible con valores propios $1/(\lambda_i - \rho)$ y sendos vectores propios correspondientes $\vec{v}_i$, $i = 1, 2, \ldots, n$. Puesto que $0 < |\lambda_1 - \rho| \ll |\lambda_i - \rho|$ para todo $i \geq 2$, se desprende que $1/(\lambda_1 - \rho)$ es el valor propio dominante de la matriz $(A - \rho I_n)^{-1}$. Entonces, podemos aplicar el método de la potencia para encontrar una aproximación a un múltiplo escalar del vector $\vec{v}_1$. Con esto hemos probado el siguiente teorema.

Teorema 8.12 **(Método de iteración inversa)** *Sea $A \in \mathfrak{M}_n(\mathbb{R})$ una matriz diagonalizable con valores propios λ_i y vectores propios correspondientes $\vec{v}_i$, $i = 1, 2, \ldots, n$. Si λ_1 tiene multiplicidad algebraica uno, $\rho \in \mathbb{R}$ es una muy buena aproximación de λ_1; esto es, $0 < |\lambda_1 - \rho| \ll |\lambda_i - \rho|$ para todo $i \geq 2$, y $\vec{u}_0 = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$ es cualquier vector tal que $\alpha_1 \neq 0$, entonces la sucesión $(\vec{u}_k)$ definida por la relación de recurrencia*

$$\left.\begin{aligned} (A - \rho I_n)\vec{z}_k &= \vec{u}_{k-1} \\ \vec{u}_k &= \frac{1}{\text{máxc}(\vec{z}_k)}\vec{z}_k \end{aligned}\right\} \tag{8.71}$$

donde z_k es la única solución del sistema $(A - \rho I_n)\vec{z}_k = \vec{u}_{k-1}$, $k = 1, 2, \ldots$, converge a un múltiplo escalar no nulo de $\vec{v}_1$; es decir, a un vector propio correspondiente a λ_1.

❍ **Nota 8.7** 1. La hipótesis de que λ_1 tenga multiplicidad uno implica que λ_1 debe ser real.

2. La suposición de que $0 < |\lambda_1 - \rho| \ll |\lambda_i - \rho|$ para todo $i \geq 2$ se hace únicamente en aras de obtener una convergencia rápida. En realidad, para la demostración del teorema basta que $0 < |\lambda_1 - \rho| < |\lambda_i - \rho|\ \forall i \geq 2$.
3. Es recomendable, para mantener estabilidad numérica, utilizar el método de Gauss con la estrategia de pivote parcial (cfr. apartado 8.2.4, pág. 838) al resolver los sistemas $(A - \rho I_n)\vec{z}_k = \vec{u}_{k-1}$.

▶ **Ejemplo 8.62** En el ejemplo 8.59 se utilizó el método de deflación de Householder para hallar los valores característicos de la matriz

$$A = \begin{bmatrix} -2 & -1 & -3 \\ -7 & 4 & -9 \\ -3 & 3 & -4 \end{bmatrix}$$

y resultó $\lambda_1 = -3$, $\lambda_2 = 2$ y $\lambda_3 = -1$. Aplicar el ciclo iterativo (8.71) del método de iteración inversa para hallar un vector propio correspondiente a $\lambda_3 = -1$.◀

Solución Utilicemos $\rho = -0.98$, que es una "muy buena aproximación" al valor propio $\lambda_3 = -1$. Para hallar las soluciones de los sistemas $(A - \rho I_n)\vec{z}_k = \vec{u}_{k-1}$ vamos a utilizar la función `pivote_parcial` que construimos en la subsección 8.2.4 (cfr. figura 8-6, pág. 842) con vector inicial $\vec{u}_0 = (1, 1, 1)$. Dado que $\rho = -0.98$,

$$A - \rho I_3 = \begin{bmatrix} -1.02 & -1.00 & -3.00 \\ -7.00 & 4.98 & -9.00 \\ -3.00 & 3.00 & -3.02 \end{bmatrix}$$

$k = 1$: Al evaluar la función `pivote_parcial` en la matriz ampliada $[\ A - \rho I_3 \quad \vec{u}_0\ ]$ se obtiene

$$\vec{z}_1 = (198.3554, 97.6842, -100.3356)$$

y, por tanto,

$$\vec{u}_1 = \frac{1}{198.3554}\vec{z}_1 = (1.0000, 0.4925, -0.5058)$$

$k = 2$: Al evaluar la función `pivote_parcial` en la matriz ampliada $[\ A - \rho I_3 \quad \vec{u}_1\ ]$ se obtiene

$$\vec{z}_2 = (50.4091, -25.2029.25.2067)$$

y, entonces,

$$\vec{u}_2 = \frac{1}{50.4091}\vec{u}_1 = (1.0000, 0.5000, -0.5000)$$

$k = 3$: Si repetimos el proceso ahora con la matriz ampliada $[\ A - \rho I_3 \quad \vec{u}_2\ ]$, se vuelve a obtener (en MATLAB con cuatro cifras significativas)

$$\vec{u}_3 = (1.0000, 0.5000, -0.5000)$$

El lector puede fácilmente verificar que efectivamente $\vec{v}_3 = \vec{u}_3$ es un vector propio correspondiente a $\lambda_3 = -1$. ✔

```
function [v, p]=iteracion_inversa(A,ro,eps,m)
%***Método de iteración inversa***
%[v,p]=iteracion_inversa(A,ro,eps,m)
%aproxima un vector propio asociado a
%un valor propio  de la matriz A. El valor
%ro es una buena aproximación al valor propio,
%eps es el margen de toleracia y m el número
%de iteraciones permisible. En la salida v es
%la aproximación del vector propio y p el
%menor número de iteraciones necesarias dentro
%del margen de tolerancia eps.
n=length(A);
u0=ones(n,1);
B=A-ro*eye(n);
B1=[B u0];
d=1;
k=0;
while (d>=eps)&(k<=m)
   k=k+1;
   z=pivote_parcial(B1)';
   [w j]=max(abs(z));
   u=z/z(j);
   d=norm(u-u0,inf);
   u0=u;
   B1=[B u];
end
p=k;
v=u;
```

Figura 8-25 • Código en MATLAB de la función `iteracion_inversa`.

Programa en MATLAB para el método de iteración inversa

La figura 8-25 contiene el código en MATLAB de la función `iteracion_inversa` para aproximar un vector propio correspondiente a un valor propio λ de una matriz A que cumplen con las condiciones del teorema 8.12. Este programa utiliza la función `pivote_parcial` de la figura 8-6 (cfr. pág. 842) para hallar las soluciones de los sistemas $(A-\rho I_n)\vec{z}_k = \vec{u}_{k-1}$ del ciclo (8.71) mediante el método de Gauss con la estrategia de pivote parcial del apartado 8.2.4. Las entradas del programa son la matriz A, una "muy buena aproximación" (`ro`) del valor propio λ, para el que se va a aproximar un vector propio correspondiente, y el margen de tolerancia (`eps`) para las iteraciones sucesivas relativas a la norma $\|\cdot\|_\infty$; la salida es la primera iteración `u_k` que satisface el margen de tolerancia.

► **Ejemplo 8.63** Utilizar la función `iteracion_inversa` para calcular un vector propio correspondiente al valor propio $\lambda_2 = 2$ de la matriz del ejemplo anterior.◄

Solución

```
EDU» [v k]=iteracion_inversa(A,2.01,.00001,20)
v =
   -1.0000
    1.0000
    1.0000
k =
     4
EDU»
```
✔

8.5.4 Método QR

Sean A una matriz real de orden n y Q_0, R_0 las matrices que se obtienen por medio de la factorización QR del método de Householder del apartado 8.4.2 (cfr. pág. 908) aplicado a la matriz A; esto es, $A = Q_0R_0$ donde Q_0 es producto de matrices de Householder —por tanto Q_0 es ortogonal (cfr. nota al pie de la página 909)— y R_0 es una matriz triangular superior. Sea $A_1 = R_0Q_0$, entonces $Q_0A_1Q_0^{-1} = Q_0R_0 = A$; luego A y A_1 tienen los mismos valores característicos por ser matrices similares. Apliquemos ahora el método de factorización QR de Householder a la matriz A_1 para obtener $A_1 = Q_1R_1$ y sea $A_2 = R_1Q_1$, entonces, por lo que acabamos de ver, A_1 y A_2 tienen los mismos valores propios y, por tanto, A y A_2

tienen los mismos valores propios. Con esto hemos hecho plausible la demostración del siguiente teorema; la cual se puede hacer fácilmente por inducción y los detalles de ésta se dejan de ejercicio al lector.

Teorema 8.13 *Sea $A \in \mathfrak{M}_n(\mathbb{R})$, y sean las matrices, definidas por recurrencia,*

$$\left.\begin{array}{ll} A & = A_0 \\ A_{k-1} & = Q_{k-1}R_{k-1} \\ A_k & = R_{k-1}Q_{k-1} \end{array}\right\} \tag{8.72}$$

donde $Q_{k-1}R_{k-1}$ es la factorización QR del método de Householder para la matriz A_{k-1}, $k = 1,2,\ldots$ Entonces A y A_k tienen los mismos valores propios para todo $k = 1,2,\ldots$

Bajo ciertas condiciones generales[16] se puede demostrar que la sucesión (A_k), definida por el ciclo iterativo (8.72) del teorema anterior, converge a una matriz diagonal por bloques; esto es, a una matriz de la forma

$$\begin{bmatrix} B_1 & * & * & \cdots & * \\ 0 & B_2 & * & \cdots & * \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & * \\ 0 & 0 & \cdots & 0 & B_m \end{bmatrix} \tag{8.73}$$

donde cada bloque B_j es una matriz 1×1 o una matriz 2×2 (los asteriscos representan componentes escalares, no necesariamente iguales entre sí). Cada bloque 2×2 corresponde a una matriz con dos valores propios complejos conjugados de la matriz A y los bloques 1×1 son valores propios reales de la matriz A. Además, si la matriz A es simétrica, cada una de las iteraciones A_k también es una matriz simétrica y la sucesión (A_k) converge a una matriz diagonal cuyos elementos en la diagonal son los valores propios de la matriz A. Por este hecho, el ciclo iterarivo (8.72) da lugar a uno de los mejores algoritmos para hallar los valores propios de una matriz; al proceso se le llama **método QR**.

En los siguientes ejemplos utilizaremos la función `ciclo_QR` de la figura 8-26 para realizar los cálculos en MATLAB del ciclo iterativo definido por las relaciones de recurrencia (8.72); en esta función se utiliza el programa `fac_QR_householder` de la figura 8-18, pág. 912, para realizar factorizaciones QR con el método de Householder.

```
function P=ciclo_QR(A,m)
%*** ciclo_QR(A,m) ***
%Calcula la iteración A_m
%del método QR
B=A;
for j=1:m
   [Q R]=fac_QR_householder(B);
   B=R*Q;
end
P=B;
```

Figura 8-26 • Función `ciclo_QR`.

▶ **Ejemplo 8.64** Sea

$$A = \begin{bmatrix} 32 & -36 & 70 & 45 \\ 5 & -5 & 10 & 5 \\ -9 & 10 & -19 & -17 \\ -1 & 2 & -4 & 4 \end{bmatrix}.$$

[16] Cfr. Capítulo 8 de la referencia bibliográfica [24].

Para calcular la primera iteración del ciclo iterativo 8.72, necesitamos encontrar la factorización QR de la matriz $A_0 = A$ con el método de Householder. Para este fin, al utilizar en MATLAB la función `fac_QR_householder` (cfr. figura 8-18, pág. 912), se obtiene

$$Q_0 = \begin{bmatrix} -0.9515 & 0.0886 & -0.1379 & 0.2603 \\ -0.1487 & -0.5660 & -0.5170 & -0.6247 \\ 0.2676 & 0.0911 & -0.8057 & 0.5206 \\ 0.0297 & -0.8146 & 0.2542 & 0.5206 \end{bmatrix} \text{ y}$$

$$R_0 = \begin{bmatrix} -33.6303 & 37.7338 & -73.2969 & -47.9924 \\ 0.0000 & -1.0779 & 2.0696 & -3.6495 \\ 0.0000 & 0.0000 & -0.5299 & 5.9240 \\ 0.0000 & 0.0000 & 0.0000 & 1.8220 \end{bmatrix}.$$

Por tanto,

$$A_1 = R_0Q_0 = \begin{bmatrix} 5.3475 & 8.0816 & 31.9812 & -95.4664 \\ 0.6056 & 3.7712 & -2.0378 & -0.1491 \\ 0.0343 & -4.8737 & 1.9328 & 2.8080 \\ 0.0542 & -1.4841 & 0.4631 & 0.9485 \end{bmatrix}$$

Podemos calcular directamente en MATLAB las iteraciones A_k mediante la función `ciclo_QR` (cfr. figura 8-26). Evaluando ésta en la matriz A con los valores $m = 20, 30, 50$ y 100, tenemos entonces

$$A_{20} = \begin{bmatrix} 6.9984 & -3.8861 & 7.3703 & -57.7992 \\ -0.0036 & 0.6037 & -1.5341 & -26.6116 \\ 0.0022 & 0.7306 & -0.0909 & 29.5848 \\ 0.0005 & -0.0394 & 0.0446 & 4.4888 \end{bmatrix},$$

$$A_{30} = \begin{bmatrix} 6.9999 & 0.7429 & -1.2448 & -58.3000 \\ 0.0001 & 5.0001 & -2.9695 & 38.1003 \\ 0.0000 & 0.0000 & 0.3275 & -11.4161 \\ 0.0000 & 0.0000 & 0.0970 & -0.3275 \end{bmatrix},$$

$$A_{50} = \begin{bmatrix} 7.0000 & 0.7428 & -1.2450 & -58.2972 \\ 0.0000 & 5.0000 & -2.9695 & 38.1046 \\ 0.0000 & 0.0000 & 0.3275 & -11.4161 \\ 0.0000 & 0.0000 & 0.0970 & -0.3275 \end{bmatrix},$$

$$A_{100} = \begin{bmatrix} 7.0000 & 0.7428 & -1.2450 & -58.2972 \\ 0.0000 & 5.0000 & -2.9695 & 38.1046 \\ 0.0000 & 0.0000 & 0.3275 & -11.4161 \\ 0.0000 & 0.0000 & 0.0970 & -0.3275 \end{bmatrix}.$$

De donde se infiere que la sucesión (A_k) converge a una matriz de la forma (8.73), con

$$B_1 = [7],\ B_2 = [5] \text{ y } B_3 = \begin{bmatrix} 0.3275 & -11.4161 \\ 0.0970 & -0.3275 \end{bmatrix}.$$

El polinomio caractéristico de la matriz B_3 está dado por

$$\lambda^2 - \text{tra}(B_2)\lambda + \det(B_2) = \lambda^2 + 1;$$

así que los valores propios de A son $\lambda_1 = 7$, $\lambda_2 = 5$, $\lambda_3 = i$ y $\lambda_4 = -i$.◄

▶ **Ejemplo 8.65** Al evaluar la función `ciclo_QR` en la matriz simétrica

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

en MATLAB resulta:

$$A_{10} = \begin{bmatrix} 2.0000 & 0.0000 & 0.0021 \\ 0.0000 & 2.0000 & 0.0036 \\ 0.0021 & 0.0036 & -1.0000 \end{bmatrix}, \quad A_{16} = \begin{bmatrix} 2.0000 & 0.0000 & 0.0000 \\ 0.0000 & 2.0000 & 0.0001 \\ 0.0000 & 0.0001 & -1.0000 \end{bmatrix},$$

$$A_{50} = \begin{bmatrix} 2.0000 & 0.0000 & 0.0000 \\ 0.0000 & 2.0000 & 0.0000 \\ 0.0000 & 0.0000 & -1.0000 \end{bmatrix}, \quad A_{100} = \begin{bmatrix} 2.0000 & 0.0000 & 0.0000 \\ 0.0000 & 2.0000 & 0.0000 \\ 0.0000 & 0.0000 & -1.0000 \end{bmatrix}.$$

Se infiere ahora que la sucesión (A_k), definida por las relaciones de recurrencia (8.72) del método QR, converge a la matriz diagonal A_{100}; luego, los valores propios de la matriz A son, $\lambda_1 = 2$, $\lambda_2 = 2$ y $\lambda_3 = -1$. De hecho, el lector puede comprobar que los valores propios exactos de la mariz A son precisamente éstos.◀

En general, *si los valores propios de la matriz A son reales y distintos entre sí o la matriz es simétrica*, al aplicar el método QR, para k grande, la matriz A_k tendrá la forma

$$\begin{bmatrix} \lambda_1 & b_2 & \cdots & b_n \\ 0 & & & \\ \vdots & & \Gamma & \\ 0 & & & \end{bmatrix} \tag{8.74}$$

Luego, λ_1 es valor propio de la matriz A y los restantes valores propios son los valores propios de la submatriz Γ; entonces, en la práctica, se debe aplicar ahora el ciclo iterativo (8.72) a la matriz Γ para obtener, a la larga, una matriz con la forma

$$\begin{bmatrix} \lambda_2 & b_3' & \cdots & b_n' \\ 0 & & & \\ \vdots & & \Gamma' & \\ 0 & & & \end{bmatrix}$$

El proceso puede repetirse hasta encontrar todos los valores propios de la matriz A.

▶ **Ejemplo 8.66** Consideremos la matriz

$$A = \begin{bmatrix} 1 & 3 & 1 \\ 4 & 2 & 3 \\ 1 & 2 & 5 \end{bmatrix}$$

Al aplicar la función `ciclo_QR` a esta matriz, se obtiene, que por primera vez en $m = 11$, la matriz

$$A_{11} = \begin{bmatrix} 7.5448 & -1.1881 & 0.7201 \\ 0.0000 & 1.0444 & 2.3313 \\ 0.0000 & 2.0670 & -0.5892 \end{bmatrix}$$

tiene la forma (8.74). Entonces

$$\Gamma = \begin{bmatrix} 1.0444 & 2.3313 \\ 2.0670 & -0.5892 \end{bmatrix}.$$

Ahora aplicamos la función `ciclo_QR` para algunos valores "grandes" de m a la matriz Γ:

$$A_{20} = \begin{bmatrix} 2.5653 & 0.3286 \\ 0.0644 & -2.1101 \end{bmatrix}, A_{60} = \begin{bmatrix} 2.5698 & 0.2643 \\ 0.0000 & -2.1146 \end{bmatrix}, A_{70} = \begin{bmatrix} 2.5698 & 0.2643 \\ 0.0000 & -2.1146 \end{bmatrix}$$

Por tanto, los valores propios de la matriz A son, aproximadamente, $\lambda_1 = 7.5448$, $\lambda_2 = 2.5698$ y $\lambda_3 = -2.1146$.◄

Programa en MATLAB para el método QR

La figura 8-28 contiene el código de la función `metodo_QR` para calcular en MATLAB los valores propios de una matriz empleando el método QR; el programa usa internamente la función `algoritmoqrB2` (figura 8-27) que, utilizando el procedimiento `fac_QR_householder`, aplica el ciclo iterativo (8.72) para obtener, a partir de una matriz dada, una matriz similar de la forma (8.74). La función emite un mensaje si, dentro de un máximo de iteraciones permitidas por el usuario, en alguna parte del ciclo no se ha logrado obtener la forma (8.74) y se tiene que aumentar el número permisible de iteraciones o considerar la posibilidad de que no haya convergencia.

```
function [P, Lambda]=algoritmoqrB2(A,m)
B=A;
n=length(A);
K1=B(2:n,1);
r=0;
while (max(abs(K1))>.000001)&(r<=m)
   [Q R]=fac_QR_householder(B);
   B=R*Q;
   K1=B(2:n,1);
   r=r+1;
end
if r>=m
   P=[];
   Lambda=[];
else
   P=B;
   Lambda=B(1,1);
end
```

Figura 8-27 • Función `algoritmoqrB2`.

► **Ejemplo 8.67** Utilizar en MATLAB la función `metodo_QR` para calcular los valores propios de la matriz

$$W = \begin{bmatrix} 10 & -14 & 10 & 7 \\ 7 & -11 & 14 & 7 \\ -2 & 2 & 1 & -4 \\ 4 & -4 & 8 & 11 \end{bmatrix}$$ ◄

Solución

```
EDU» W=[10 -14 10 7;7 -11 14 7;-2 2 1 -4;4 -4 8 11];
EDU» metodo_QR(W,100)

Proceso llevado a cabo con éxito.
Valores propios:
ans =
    7.0000
   -4.0000
    5.0000
    3.0000
EDU»
```
✔

```
function Lambdas=metodo_QR(A,m)
%*** Método QR ***
%Lambdas=metodo_QR(A,m)
%calcula los valores propios de
%la matriz A mediante el método QR
%dentro un número máximo permisible
%de iteraciones m.
n=length(A);
A1=A;
for j=1:n-1
   [P lambda]=algoritmoqrB2(A1,m);
   if isempty(lambda)==1
      L1=[];
      break
   else
      L(j)=lambda;
      P(:,1)=[];
      P(1,:)=[];
      A1=P;
      L1=L;
   end
end
if isempty(L1)==1
   fprintf('\nMáximo número de iteraciones excedido: \%3d.\n',m)
   fprintf('Incrementar el número de iteraciones o\n')
   fprintf('posiblemente no hay convergencia.')
   Lambdas=[];
else
   fprintf('\nProceso llevado a cabo con éxito.\n')
   fprintf('Valores propios:\n')
   Lambdas=[L1 A1];
end
```

Figura 8-28 • Función metodo_QR.

▶ **Ejemplo 8.68** Repetir el ejemplo precedente con la matriz $G = \begin{bmatrix} 7 & 4 & -4 & 7 \\ 4 & 11 & -9 & 1 \\ -4 & -9 & 13 & 1 \\ 7 & 1 & 1 & 12 \end{bmatrix}$ ◀

Solución

```
EDU» G=[7 4 -4 7;4 11 -9 1;-4 -9 13 1;7 1 1 12];
EDU» metodo_QR(G,100)

Proceso llevado a cabo con éxito.
Valores propios:
ans =
   23.6031
   15.6419
    2.9202
    0.8348
EDU»
```
✔

Puede ser que el método QR no converja a una matriz en la forma (8.73); hacemos patente este hecho en el siguiente ejemplo.

▶ **Ejemplo 8.69** Sea

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & -2 \\ -1 & 0 & 1 \end{bmatrix}.$$

Si utilizamos en MATLAB la función `fac_QR_householder` para encontrar la factorización QR de esta matriz obtenemos

$$Q_0 = \begin{bmatrix} -0.5774 & -0.4082 & -0.7071 \\ 0.5774 & -0.8165 & 0.0000 \\ -0.5774 & -0.4082 & 0.7071 \end{bmatrix} \quad \text{y} \quad R_0 = \begin{bmatrix} 1.7321 & -0.5774 & -1.7321 \\ 0.0000 & 0.8165 & 1.2247 \\ 0.0000 & 0.0000 & 0.7071 \end{bmatrix},$$

entonces,

$$A_1 = R_0Q_0 \begin{bmatrix} -0.3333 & 0.4714 & -2.4495 \\ -0.2357 & -1.1667 & 0.8660 \\ -0.4082 & -0.2887 & 0.5000 \end{bmatrix}.$$

Al evaluar ahora `fac_QR_householder` en A_1, resulta

$$Q_1 = \begin{bmatrix} -0.5774 & 0.5774 & -0.5774 \\ -0.4082 & -0.8165 & -0.4082 \\ -0.7071 & 0.0000 & 0.7071 \end{bmatrix} \quad \text{y} \quad R_1 = \begin{bmatrix} 0.5774 & 0.4082 & 0.7071 \\ 0.0000 & 1.2247 & -2.1213 \\ 0.0000 & 0.0000 & 1.4142 \end{bmatrix},$$

y, por tanto,

$$A_2 = R_1Q_1 \begin{bmatrix} -1.0000 & 0.0000 & 0.0000 \\ 1.0000 & -1.0000 & -2.0000 \\ -1.0000 & 0.0000 & 1.0000 \end{bmatrix}.$$

que es nuevamente la matriz A, la cual no está en la forma (8.74).◀

De esta manera, si se aplica la función `metodo_QR` a la matriz A del ejemplo precedente, el programa emitirá el mensaje de que se debe aumentar el número de iteraciones permisibles o de que posiblemente no hay convergencia (cfr. figura 8-28). Sin embargo, calcular los valores propios de la matriz A es muy simple y se puede hacer manualmente (pruebe con la matriz A^t). Lo mismo ocurrirá para la matriz del ejemplo 8.64; porque, como vimos en éste, la sucesión (A_k) converge a una matriz que está en la forma (8.74) y los elementos en la cuarta columna de las iteraciones serán distintos de cero bajo la diagonal. Por esta razón el programa `metodo_QR`, como es natural, es limitado y conviene —cuando emita el mencionado mensaje— utilizarlo en combinación con la función `ciclo_QR` para ir observando el comportamiento de las sucesiones y actuar, cuando sea posible, como se hizo en el ejemplo 8.64. En general, el programa `metodo_QR` dará resultados correctos cuando actúe sobre matrices simétricas, o que no lo son, pero tienen valores propios reales distintos entre sí.

Cálculo de valores propios con reducción de Hessenberg y método QR

Todos los métodos para estimar numéricamente valores propios son iterativos. Los procesos para efectuar cada iteración pueden requerir de una enorme cantidad de operaciones que en la práctica es inviable ejecutar. Una manera de atacar el problema es trabajar con matrices similares a la matriz A que tengan un formato que implique menos operaciones al implementar el algoritmo correspondiente. La reducción a la forma superior de Hessenberg, que vimos en el apartado 8.4.3 (pág. 913), es muy útil para ese fin. Una ventaja sumamente importante de reducir a forma superior de Hessenberg cuando la matriz es simétrica, es que la reducción resulta una matriz tridiagonal simétrica (cfr. teorema 8.10, pág. 915). Este procedimiento, como antes, funcionará generalmente cuando se aplique a matrices con valores propios reales distintos entre sí o a matrices simétricas. Hacemos patente a continuación este método.

Algoritmo 8.11 (Método QR con reducción de Hessenberg)

1. *La matriz A se reduce a su forma superior de Hessenberg A_H.*
2. *Se aplica el algorimo QR a la matriz A_H hasta que A_{H_k} tenga la forma (8.74).*
3. *La matriz Γ se reduce a su forma superior de Hessenberg Γ_H.*
4. *Ahora se aplica el método QR a la matriz Γ_H hasta obtener una matriz de la forma (8.74), pero de orden $n-1\times n-1$.*
5. *Se repite el proceso hasta que se obtiene una matriz en forma diagonal con los valores propios en la diagonal.*

► **Ejemplo 8.70** Consideremos la matriz

$$A=\begin{bmatrix}1&3&2\\5&3&2\\2&1&7\end{bmatrix}.$$

Entonces, al evaluar la función reduccion_hhh (cfr. figura 8-19, pág. 916) en la matriz A, obtenemos la matriz similar superior de Hessenberg

$$A_H=\begin{bmatrix}1.0000&-3.5282&0.7428\\-5.3852&4.5862&-2.9655\\0&-1.9655&5.4138\end{bmatrix}.$$

Ahora apliquemos el algoritmo QR a la matriz A_H. Utilizando la función `ciclo_QR` (cfr. figura 8-26, pág. 940) en MATLAB resulta, por primera vez en $m=15$,

$$A_{H_{15}}=\begin{bmatrix}8.9257&-1.0270&-1.6995\\0.0000&4.1339&-1.0289\\0.0000&-0.0008&-2.0596\end{bmatrix}$$

y, por tanto, $\lambda_1=8.9257$. Ahora se repite el procedimiento con la submatriz

$$B=\begin{bmatrix}4.1339&-1.0289\\-0.0008&-2.0596\end{bmatrix}.$$

No es necesario reducir la matriz B a forma superior de Hessenberg pues ya está en ese formato; entonces aplicamos nuevamente la función ciclo_QR a la matriz B, para obtener, por primera vez

$$B_4=\begin{bmatrix}4.1340&-1.0281\\0.0000&-2.0597\end{bmatrix}.$$

Así, los valores propios de la matriz A son $\lambda_1=8.9257$, $\lambda_2=4.1340$ y $\lambda_3=-2.0597$.◄

8.5.5 Método QR con reducción de Hessenberg y desplazamientos

El método QR, aun con reducción de Hessenberg, puede tener una convergencia muy lenta si la matriz posee valores propios cercanos entre sí. Para lograr acelerar notablemente la convergencia de las iteraciones planteamos a continuación uno de los métodos más utilizados para este fin, el método QR con reducción de Hessenberg y desplazamientos.

Supongamos que una matriz real A de orden n tiene valores propios distintos entre sí $\lambda_1, \lambda_2, \ldots, \lambda_n$, y que ρ es una muy buena aproximación de λ_n. Entonces, como vimos antes en el método de iteración inversa, la matriz $A - \rho I_n$ tiene los valores propios $\lambda_1 - \rho$, $\lambda_2 - \rho, \ldots, \lambda_n - \rho$. Apliquemos el método QR con reducción de Hessenberg a la matriz $B = A - \rho I_n$; si representamos por $B_k = [b_{ij}^{(k)}]$ la k-ésima iteración de este método, entonces, de manera bastante rápida, la componente $b_{n,n-1}^{(k)}$ tenderá a cero y, por ende,

$$B_k + \rho I_n = \begin{bmatrix} \widehat{B}_k & * \\ 0 & b_{nn}^{(k)} + \rho \end{bmatrix}.$$

Por construcción B_k es similar a $B = A - \rho I_n$, luego $B_k + \rho I_n$ es similar a la matriz A y tienen los mismos valores propios; así que $\lambda_n = b_{nn}^{(k)} + \rho$. Los restantes valores propios de A son los valores propios de la matriz $\widehat{B}_k$ y se puede aplicar el algoritmo QR con reducción de Hessenberg a ésta para hallarlos. En la práctica difícilmente se puede disponer de una buena aproximación para λ_n, entonces, en cada parte del ciclo, se toma ρ igual a la última entrada de la diagonal de la iteración. Resumimos este importante método en el siguiente algoritmo.

Algoritmo 8.12 (Método QR con reducción de Hessenberg y desplazamientos)

1. *Sea A_0 la forma superior de Hessenberg de la matriz A.*
2. *Se define por recurrencia, para cada $k = 1, 2, \ldots$, la siguiente sucesión de matrices $A_k = [a_{ij}^{(k)}]$:*

   (a) $\rho_{k-1} = a_{nn}^{(k-1)}$, la última componente de la diagonal de la matriz A_{k-1}.

   (b) $A_{k-1} - \rho_{k-1} I_n = Q_{k-1} R_{k-1}$ (factorización QR de A_{k-1}).

   (c) $A_k = R_{k-1} Q_{k-1} + \rho_{k-1} I_n$.

3. *Se efectúa el ciclo iterativo definido en el inciso anterior hasta que A_k tiene la forma*

$$A_k = \begin{bmatrix} \widehat{B}_k & * \\ 0 & a_{nn}^{(k)} + \rho_{k-1} \end{bmatrix}.$$

   La última componente de la diagonal de A_k es el valor propio λ_n de la matriz A.

4. *Se repiten los pasos 1, 2 y 3 con la matriz $\widehat{B}_k$ para obtener el valor propio λ_{n-1}; y así sucesivamente, hasta obtener todos los valores propios de la matriz A.*

Es fácil probar (se deja de ejercicio al lector) que para todo k las matrices A_k y A_{k-1} son ortogonalmente similares.

► **Ejemplo 8.71** Calcular los valores propios de la matriz

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 5 & 3 & 2 \\ 2 & 1 & 7 \end{bmatrix}$$

del ejemplo 8.70.◄

Solución Si utilizamos la función `reduccion_hhh` de la figura 8-19, pág. 916, para calcular la reducción a forma superior de Hessenberg de la matriz A obtenemos

$$A_0 = \begin{bmatrix} 1.0000 & -3.5282 & 0.7428 \\ -5.3852 & 4.5862 & -2.9655 \\ 0 & -1.9655 & 5.4138 \end{bmatrix}.$$

$k = 1$: $\rho_0 = a_{33}^{(0)} = 5.4138$ y al calcular la factorización QR de la matriz $A - \rho_0 I_3$, con la función `fac_QR_householder` de la figura 8-18, pág. 912, resulta

$$Q_0 = \begin{bmatrix} -0.6339 & 0.5772 & -0.5147 \\ -0.7734 & -0.4731 & 0.4219 \\ 0 & 0.6655 & 0.7464 \end{bmatrix} \quad \text{y} \quad R_0 = \begin{bmatrix} 6.9629 & 2.8766 & 1.8227 \\ 0.0000 & -2.9532 & 1.8318 \\ 0.0000 & 0.0000 & -1.6335 \end{bmatrix}$$

y, por tanto,

$$A_1 = R_0 Q_0 + \rho_0 I_3 = \begin{bmatrix} -1.2248 & 3.8714 & -1.0101 \\ 2.2841 & 8.0302 & 0.1212 \\ 0.0000 & -1.0872 & 4.1946 \end{bmatrix}.$$

$k = 2$: $\rho_1 = a_{33}^{(1)} = 4.1946$ y $A_1 - \rho_1 I_3 = Q_1 R_1$ con

$$Q_1 = \begin{bmatrix} -0.9215 & -0.3796 & 0.0819 \\ 0.3884 & -0.9008 & 0.1944 \\ 0.0000 & 0.2109 & 0.9775 \end{bmatrix} \quad \text{y} \quad R_1 = \begin{bmatrix} 5.8811 & -2.0779 & 0.9779 \\ 0.0000 & -5.1539 & 0.2743 \\ 0.0000 & 0 & -0.0592 \end{bmatrix}$$

así que

$$A_2 = R_1 Q_1 + \rho_1 I_3 = \begin{bmatrix} -2.0318 & -0.1547 & 1.0338 \\ -2.0017 & 8.8950 & -0.7337 \\ 0.0000 & -0.0125 & 4.1368 \end{bmatrix}$$

$k = 3$: $\rho_2 = a_{33}^{(2)} = 4.1368$ y $A_2 - \rho_2 I_3 = Q_2 R_2$ con

$$Q_2 = \begin{bmatrix} -0.9512 & 0.3086 & -0.0008 \\ -0.3086 & -0.9512 & 0.0026 \\ 0.0000 & 0.0027 & 1.0000 \end{bmatrix} \quad \text{y} \quad R_2 = \begin{bmatrix} 6.4852 & -1.3214 & -0.7568 \\ 0.0000 & -4.5737 & 1.0170 \\ 0.0000 & 0 & -0.0028 \end{bmatrix}$$

entonces,

$$A_3 = R_2 Q_2 + \rho_2 I_3 = \begin{bmatrix} -1.6239 & 3.2565 & -0.7657 \\ 1.4117 & 8.4899 & 1.0051 \\ 0.0000 & 0.0000 & 4.1340 \end{bmatrix}.$$

Por tanto, $\lambda_3 = 4.1340$ y aplicamos el algoritmo a la submatriz

$$\widehat{B}_3 = \begin{bmatrix} -1.6239 & 3.2565 \\ 1.4117 & 8.4899 \end{bmatrix}.$$

$A_0 = \widehat{B}_3$ (ya está en forma superior de Hessenberg).

$k = 1$: $\widehat{\rho}_0 = b_{33}^{(0)} = 8.4899$, $B_0 - \widehat{\rho}_0 I_2 = Q_0 R_0$, con

$$Q_0 = \begin{bmatrix} -0.9904 & 0.1382 \\ 0.1382 & 0.9904 \end{bmatrix} \quad \text{y} \quad R_0 \begin{bmatrix} 10.2119 & -3.2252 \\ 0 & 0.4502 \end{bmatrix}$$

y, por tanto,

$$B_1 = R_0 Q_0 + \widehat{\rho}_0 I_2 = \begin{bmatrix} -2.0698 & -1.7826 \\ 0.0622 & 8.9358 \end{bmatrix}.$$

$k = 2$: $\widehat{\rho}_1 = b_{22}^{(1)} = 8.9358$, $B_1 - \widehat{\rho}_1 I_2 = Q_1 R_1$, con

$$Q_1 = \begin{bmatrix} -1.0000 & 0.0057 \\ 0.0057 & 1.0000 \end{bmatrix} \quad \text{y} \quad R_1 \begin{bmatrix} 11.0057 & 1.7826 \\ 0.0000 & -0.0101 \end{bmatrix}$$

y, por ende,

$$B_2 = R_1 Q_1 + \widehat{\rho}_1 I_2 = \begin{bmatrix} -2.0597 & 1.8448 \\ -0.0001 & 8.9257 \end{bmatrix}.$$

$k = 3$: $\widehat{\rho}_2 = b_{22}^{(2)} = 8.9257$, $B_2 - \widehat{\rho}_2 I_2 = Q_2 R_2$, con

$$Q_2 = \begin{bmatrix} -1.0000 & 0.0000 \\ 0.0000 & 1.0000 \end{bmatrix} \quad \text{y} \quad R_2 \begin{bmatrix} 10.9854 & -1.8448 \\ 0.0000 & 0.0000 \end{bmatrix}$$

y, entonces,

$$B_3 = R_2 Q_2 + \widehat{\rho}_2 I_2 = \begin{bmatrix} -2.0597 & -1.8448 \\ 0.0000 & 8.9257 \end{bmatrix}.$$

Luego, los valores propios de la matriz A son $\lambda_3 = 4.1340$, $\lambda_2 = -2.0597$ y $\lambda_1 = 8.9257$. ✔

Ahora podemos comparar las 19 iteraciones realizadas en el ejemplo 8.70 utilizando el *método QR con reducción de Hessenberg*, contra las 7 iteraciones efectuadas mediante el *método QR con reducción de Hessenberg y desplazamientos* que efectuamos en el ejemplo precedente para hallar los valores propios de la misma matriz.

Programa en MATLAB para el método QR con reducción de Hessenberg y desplazamientos

La figura 8-29 contiene el código de la función `metQR_des` para calcular en MATLAB valores propios de matrices por medio del método QR con redución de Hessenberg y desplazamientos. El programa utiliza internamente la función `desplazamientos1` que es el código en MATLAB, contenido en la figura 8-30, del ciclo de recurrencia de los pasos 1 a 3 del algoritmo 8.12.

```
function lambdas=metQR_des(A)
%metQR_des(A) calcula los valores
%propios de la matriz A por el
%método QR con reducción de
%Hessenbereg y desplazamientos.
m=length(A);
J=A;
for j=1:m-1
   [J L(j)]=desplazamientos1(J);
   J=J(1:m-j,1:m-j);
end
lambdas=[J(1,1); L'];
```

Figura 8-29 • Código de la función `metQR_des`.

```
function [B, lambda]=desplazamientos1(A)
n=length(A);
b=A(n,n-1);
A0=reduccion_hhh(A);
while abs(b)>.0000001
   rho=A0(n,n);
   [Q R]=fac_QR_householder(A0-rho*eye(n));
      A1=R*Q+rho*eye(n);
      rho=A1(n,n);
      b=A1(n,n-1);
      A0=A1;
   end
   B=A1;
   lambda=A1(n,n);
```

Figura 8-30 • Función `desplazamientos1`.

► **Ejemplo 8.72** Calcular, utilizando en MATLAB la función `metQR_des`, los valores propios de la matriz

$$A=\begin{bmatrix}16 & 3 & 9 & 0 & 5\\ 3 & 10 & 8 & -8 & 7\\ 9 & 8 & 12 & -12 & 8\\ 0 & -8 & -12 & 31 & -8\\ 5 & 7 & 8 & -8 & 7\end{bmatrix}. ◄$$

Solución

```
EDU» A=[16 3 9 0 5;3 10 8 -8  7;9 8 12 -12 8
0 -8 -12 31 -8;5 7 8 -8 7];
EDU» metQR_des(A)

ans =
   46.3364
    0.8700
    0.5623
    7.3570
   20.8744
EDU»
```
✔

8.5.6 Método de Jacobi para aproximar valores y vectores propios de matrices simétricas

El cálculo de valores y vectores propios de matrices simétricas con componentes reales aparece con mucha frecuencia en una gran variedad de aplicaciones. Existen varios métodos numéricos especialmente diseñados para aproximar los valores propios de una matriz simétrica; en este apartado estudiaremos uno de estos algoritmos que es clásico y de gran uso en la práctica, el método de Jacobi. Con este procedimiento, de manera simultánea, se pueden calcular los valores y vectores propios correspondientes de una matriz simétrica real. Recordemos que una matriz simétrica de orden n, con entradas reales, tiene todos sus valores propios reales y además posee un conjunto de n vectores propios ortonormales.

Supongamos que $A \in \mathfrak{M}_n(\mathbb{R})$ es una matriz simétrica. Definimos, en forma recurrente, la siguiente sucesión de matrices:

$$\Lambda_k = R_k^t \Lambda_{k-1} R_k$$

$k = 1, 2, \ldots$, donde $\Lambda_0 = A$ y R_{k-1} es una matriz de rotación —por tanto una matriz ortogonal— definida por las relaciones (8.59) del apartado 8.4.4, pág. 917, tal que el producto $R_{k-1}^t \Lambda_{k-1} R_{k-1}$ transforma una componente de mayor valor absoluto fuera de la diagonal de la matriz Λ_{k-1} en cero. Dado que la matriz A es simétrica, es fácil demostrar por inducción que las matrices Λ_k son también simétricas. Siendo así, se puede probar que

$$\lim_{k\to\infty} \Lambda_k = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \Lambda$$

donde los λ_i son los valores propios de la matriz A. Luego, para k grande,

$$\Lambda = R_k^t R_{k-1}^t \cdots R_1^t A R_1 R_2 \cdots R_k.$$

Entonces, si $R = R_1 R_2 \cdots R_k$, esta matriz, por ser producto de matrices ortogonales, es una matriz ortogonal y

$$AR = R\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

De esta manera, si $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ son las columnas de R, se tiene

$$\begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 & \cdots & A\vec{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \cdots & \lambda_n\vec{v}_n \end{bmatrix}$$

y, por tanto, cada vector $\vec{v}_i$ es un vector propio correspondiente al valor propio λ_i, $i = 1, 2, \ldots, n$.

Sean ahora $M = [\alpha_{ij}] \in \mathfrak{M}_n(\mathbb{R})$ una matriz simétrica y α_{pq} una entrada cuya magnitud es mayor que cualquiera de las magnitudes de las componentes fuera de la diagonal de la matriz M; veamos cuáles deben ser las entradas de una matriz de rotación R para que el producto $D = R^t MR$ transforme las componentes α_{pq} y α_{qp} en entradas nulas de la matriz simétrica D. Recordemos que la matriz de rotación $R = [r_{ij}]$ es la matriz identidad, excepto en las entradas $r_{pp} = c = r_{qq}$ y $r_{pq} = -s$, $r_{qp} = s$ (cfr. (8.59), pág. 921) esto es,

$$\begin{array}{ccc} & \begin{array}{ccccccc} & & \text{columna } p & & & & \\ & & \downarrow & & & & \end{array} & \\ R = & \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & c & \cdots & -s & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & s & \cdots & c & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow \text{fila } p \\ \\ \leftarrow \text{fila } q \\ \\ \\ \end{matrix} \\ & \begin{array}{ccccccc} & & & & \uparrow & & \\ & & & & \text{columna } q & & \end{array} & \end{array}$$

Así, si $B = [b_{ij}]$ y

$$B = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1p} & \cdots & \alpha_{1q} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{p1} & \cdots & \alpha_{pp} & \cdots & \alpha_{pq} & \cdots & \alpha_{pn} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \alpha_{q1} & \cdots & \alpha_{qp} & \cdots & \alpha_{qq} & \cdots & \alpha_{qn} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{nn} & \cdots & \alpha_{np} & \cdots & \alpha_{nq} & \cdots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & c & \cdots & -s & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & s & \cdots & c & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix},$$

se tiene

$$\begin{aligned} b_{jk} &= \alpha_{jk}, \text{ si } k \neq p \text{ y } k \neq q \\ b_{jp} &= \alpha_{jp}c + \alpha_{jq}s \\ b_{jq} &= -\alpha_{jp}s + \alpha_{jq}c \end{aligned}$$

para cada $j = 1, 2, \ldots, n$. Es decir, únicamente las columnas p y q de M son alteradas bajo el producto. Luego, ya que $D = R^t MR = R^t B$, es fácil ver que si $D = [d_{ij}]$, sólo las filas p y q de B son alteradas bajo el producto y

$$\begin{aligned} d_{pp} &= \alpha_{pp}c^2 + \alpha_{qq}s^2 + 2\alpha_{pq}cs \\ d_{qq} &= \alpha_{pp}s^2 + \alpha_{qq}c^2 - 2\alpha_{pq}cs \\ d_{pq} &= \alpha_{pq}(c^2 - s^2) + (\alpha_{qq} - \alpha_{pp})cs \end{aligned} \tag{8.75}$$

Las demás componentes de D se pueden hallar por simetría con las filas de la matriz B que no se alteraron con el producto. Ya que R es una matriz de rotación, $c = \cos\theta$ y $s = \operatorname{sen}\theta$; por tanto, resta sólo encontrar el ángulo de rotación adecuado para que d_{pq} sea cero. Para ello definamos

$$\begin{aligned} \omega &= \cot 2\theta \\ &= \frac{\cos 2\theta}{\operatorname{sen} 2\theta} \\ &= \frac{\cos^2\theta - \operatorname{sen}^2\theta}{2\operatorname{sen}\theta\cos\theta} \\ &= \frac{c^2 - s^2}{2cs}\,; \end{aligned}$$

esto es,

$$\omega = \frac{c^2 - s^2}{2cs} \tag{8.76}$$

Ya que la componente d_{pq} tiene que ser nula, de (8.75) se debe tener, si $\alpha_{pq} \neq 0$,

$$\frac{c^2 - s^2}{cs} = \frac{\alpha_{pp} - \alpha_{qq}}{\alpha_{pq}}$$

y, por tanto,

$$\omega = \frac{\alpha_{pp} - \alpha_{qq}}{2\alpha_{pq}}\,.$$

Luego, con

$$\theta = \frac{1}{2}\operatorname{arccot}(\omega)$$

se obtiene el efecto deseado. Sin embargo, para evitar inestabilidad numérica, mejor se define

$$\mu = \tan\theta$$

y al multiplicar por $1/c^2$ el numerador y el denominador en (8.76) se obtiene

$$\omega = \frac{1 - \dfrac{s^2}{c^2}}{2\dfrac{s}{c}} = \frac{1 - \mu^2}{2\mu}$$

que conduce a la ecuación cuadrática

$$\mu^2 + 2\omega\mu - 1 = 0.$$

Resolviendo ésta,

$$\mu = -\omega \pm \sqrt{\omega^2 + 1}.$$

Si $\omega \geq 0$, la raíz con menor módulo es

$$\mu = -\omega + \sqrt{\omega^2+1} = \frac{1}{\omega + \sqrt{\omega^2+1}}.$$

Mientras que si $\omega < 0$, la raíz con menor módulo es

$$\mu = -\omega - \sqrt{\omega^2+1} = -\frac{1}{-\omega + \sqrt{\omega^2+1}}.$$

De aquí, si sign es la función que definimos en el apartado 8.4.1 (pág. 902), la raíz de menor módulo, en cualquier caso, está dada por

$$\mu = \frac{\operatorname{sign}(\omega)}{|\omega| + \sqrt{\omega^2+1}}$$

y se puede demostrar que hay estabilidad numérica eligiendo este valor de μ. Por tanto, el efecto de transformar en cero la componente α_{pq} al realizar el producto R^tMR se obtiene con

$$c = \cos\theta = \frac{1}{\sqrt{\mu^2+1}}$$
$$s = \operatorname{sen}\theta = c\mu.$$

Resumimos la información precedente en el siguiente algoritmo.

Algoritmo 8.13 **(Método de Jacobi para calcular valores y vectores propios)** *Sea A una matriz simétrica. Para calcular los valores y vectores propios de A se procede de la siguiente manera:*

1. *Se establece $\Lambda_0 = A$.*
2. *Se define, en forma recurrente, la sucesión de matrices*

$$\Lambda_k = R_k^t \Lambda_{k-1} R_{k-1}$$

para $k = 1, 2, \ldots$ La matriz R_k es una matriz de rotación con parámetros

$$c_k = \frac{1}{\sqrt{\mu_k^2+1}}$$
$$s_k = c_k\mu_k$$

con

$$\mu_k = \frac{sign(\omega_k)}{|\omega_k| + \sqrt{\omega_k^2+1}}$$

y

$$\omega_k = \frac{\alpha_{pp}^{(k-1)} - \alpha_{qq}^{(k-1)}}{2\alpha_{pq}^{(k-1)}}$$

donde $\alpha_{pq}^{(k-1)}$ es una componente de máximo valor absoluto fuera de la diagonal de la matriz[17] *$\Lambda_{k-1} = \left[\alpha_{ij}^{(k-1)}\right]$, $k = 1, 2, \ldots$*

[17] Esto es, $|\alpha_{pq}^{(k-1)}| \geq |\alpha_{ij}^{(k-1)}| \forall i, j, i \neq j$.

3. *Las iteraciones se realizan hasta que, por primera vez, el máximo valor absoluto de las componentes fuera de la diagonal de la matriz Λ_k es menor que un valor especificado T $(0 < T \ll 1)$.*
4. *Las componentes en la diagonal de la matriz Λ_k, del inciso anterior, serán las aproximaciones para los valores propios de la matriz A y las columnas de la matriz $R = R_1R_2\cdots R_k$ serán las aproximaciones para los vectores propios correspondientes, respectivamente.*

▶ **Ejemplo 8.73** Sea

$$A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}.$$

Calcular los valores y vectores propios correspondientes de esta matriz por el método de Jacobi.◀

Solución Una componente con máximo valor absoluto fuera de la diagonal de $\Lambda_0 = A$ es $\alpha_{12}^{(0)} = -1$, entonces

$$\omega_1 = \frac{\alpha_{11}^{(0)} - \alpha_{22}^{(0)}}{2\alpha_{12}^{(0)}} = \frac{2-(2)}{2(-1)} = 0;$$

por tanto,

$$\mu_1 = \frac{\operatorname{sign}(\omega_1)}{|\omega_1| + \sqrt{\omega_1^2+1}} = \frac{1}{0+\sqrt{(0)^2+1}} = 1,$$

$$c_1 = \frac{1}{\sqrt{\mu_1^2+1}} = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2},$$

$$s_1 = c_1\mu_1 = \frac{\sqrt{2}}{2}1 = \frac{\sqrt{2}}{2}.$$

Luego

$$R_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

y, entonces,

$$\begin{aligned} \Lambda_1 &= R_1^t\Lambda_0R_1 \\ &= \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}^t \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -\sqrt{2} \\ 0 & -\sqrt{2} & 2 \end{bmatrix}. \end{aligned}$$

Una componente con módulo máximo fuera de la diagonal de Λ_1 es $\alpha_{32}^{(1)} = -\sqrt{2}$, entonces

$$\omega_2 = \frac{\alpha_{33}^{(1)} - \alpha_{22}^{(1)}}{2\alpha_{32}^{(1)}} = \frac{2-(3)}{2(-\sqrt{2})} = \frac{\sqrt{2}}{4};$$

por tanto,

$$\mu_2 = \frac{\operatorname{sign}(\omega_2)}{|\omega_2| + \sqrt{\omega_2^2 + 1}} = \frac{1}{\frac{\sqrt{2}}{4} + \sqrt{\left(\frac{\sqrt{2}}{4}\right)^2 + 1}} = \frac{\sqrt{2}}{2},$$

$$c_2 = \frac{1}{\sqrt{\mu_2^2 + 1}} = \frac{1}{\sqrt{\left(\frac{\sqrt{2}}{2}\right)^2 + 1}} = \frac{\sqrt{6}}{3},$$

$$s_2 = c_2\mu_2 = \frac{\sqrt{6}}{3}\frac{\sqrt{2}}{2} = \frac{\sqrt{3}}{3}.$$

Luego,

$$R_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \end{bmatrix}$$

y, entonces,

$$\Lambda_2 = R_2^t \Lambda_1 R_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \end{bmatrix}^t \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -\sqrt{2} \\ 0 & -\sqrt{2} & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Y ya que

$$R = R_1R_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \end{bmatrix}$$

los valores propios de A son $\lambda_1 = \lambda_3 = 1$ con vectores propios correspondientes $\vec{v}_1 = \sqrt{2}/2(1,\ 1{,}0)$, $\vec{v}_3 = \sqrt{6}/6(-1,1,2)$; y $\lambda_2 = 4$ con vector propio correspondiente $\vec{v}_2 = \sqrt{3}/3(-1,1,-1)$.

Comprobación:

$$R^{-1}AR = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \end{bmatrix}^t \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad ✔$$

Se deja como ejercicio al lector la implementación del algoritmo 8.13 en MATLAB o en algún otro lenguaje de programación.

8.6 Ejercicios propuestos

El lector encontrará la respuesta a los ejercicios en cursiva en el apéndice E al final del libro.

Aritmética de la computadora y errores de redondeo (respuestas en página 1099)

En los ejercicios 1 a 8 representar el número R en aritmética de punto flotante con n cifras significativas *(a)* mediante truncamiento y *(b)* por redondeo.

1 $R = \pi$, $n = 5$.

2 $R = \sqrt{2}$, $n = 4$.

3 $R = e$, $n = 4$.

4 $R = \sqrt[3]{5}$, $n = 5$.

5 $R = \text{sen}(\pi/8)$, $n = 4$.

6 $R = 0.9999999$, $n = 5$.

7 $R = 0.000018$, $n = 4$.

8 $R = 0.000018$, $n = 5$.

En los ejercicios 9 a 18 calcular *(a)* el error absoluto y *(b)* el error relativo que se comete al aproximar el número R con el número R^*.

9 $R = 9!$, $R^* = 365\,000$.

10 $R = 12!$, $R^* = 47\,000\,000$.

11 $R = 1/3$, $R^* = 0.333$.

12 $R = 1/6$, $R^* = 0.6666$.

13 $R = \sqrt{2}$, $R^* = 1.414$.

14 $R = \sqrt{3}$, $R^* = 1.732$.

15 $R = \pi$, $R^* = 3.1416$.

16 $R = e$, $R^* = \frac{8}{3}$.

17 $R = e$, $R^* = \frac{65}{24}$.

18 $R = \cos(1)$, $R^* = 0.5417$.

En los ejercicios 19 a 23 encontrar el mayor intervalo (a, b) en el que debe estar R^* para aproximar a R con un error absoluto inferior a ε.

19 $R = 120$, $\varepsilon = 10^{-2}$.

20 $R = 540$, $\varepsilon = 10^{-3}$.

21 $R = 76$, $\varepsilon = 10^{-3}$.

22 $R = 750$, $\varepsilon = 10^{-2}$.

23 $R = 5\,300$, $\varepsilon = 2 \times 10^{-4}$.

24 Repetir el ejercicio 19 si el error es relativo.

25 Repetir el ejercicio 21 si el error es relativo.

26 Repetir el ejercicio 23 si el error es relativo.

En los ejercicios 27 a 32 realizar la operación indicada *(a)* en forma exacta, *(b)* con aritmética de punto flotante con truncamiento a tres cifras significativas, *(c)* con aritmética de punto flotante con redondeo a tres cifras significativas.

27 $\frac{3}{7} + \frac{5}{9}$.

28 $\frac{3}{5} + \frac{1}{4}$.

29 $\frac{5}{7} \cdot \frac{7}{5}$.

30 $\frac{1}{2} + \frac{3}{5} - \frac{1}{3}$.

31 $\frac{1}{3} + \frac{2}{5} - \frac{2}{7}$.

32 $\frac{1}{7} - \frac{2}{3} + \frac{1}{5}$.

Métodos directos para resolver sistemas lineales (respuestas en páginas 1099-1100)

En los ejercicios 33 a 38 aplicar el método de Gauss —algoritmo 8.1, pág 823— para resolver el sistema lineal dado o determinar si no tiene solución única.

33
$$\begin{array}{rcrcrcrcr} x_1 &+& x_2 &-& x_3 &+& x_4 &=& 2\\ 2x_1 &-& 3x_2 &+& x_3 &-& x_4 &=& 8\\ x_1 &-& 2x_2 &+& 3x_3 &+& 2x_4 &=& 15\\ 3x_1 &-& x_2 &-& x_3 &-& 3x_4 &=& -2 \end{array}$$

34
$$\begin{array}{rcrcrcrcr} 2x_1 &-& x_2 &+& x_3 &+& 3x_4 &=& 6\\ 3x_1 &-& x_2 &+& 2x_3 &+& x_4 &=& 0\\ 2x_1 &-& 2x_2 &+& x_3 &+& 2x_4 &=& 1\\ 2x_1 &+& x_2 &+& x_3 &+& 3x_4 &=& 10 \end{array}$$

35
$$\begin{array}{rcrcrcrcr} x_1 &+& x_2 &-& x_3 &+& x_4 &=& 2\\ 2x_1 &-& 3x_2 &+& x_3 &-& x_4 &=& 8\\ -x_1 &+& 4x_2 &-& x_3 &+& 5x_4 &=& 6\\ 3x_1 &-& x_2 &-& x_3 &-& 3x_4 &=& -2 \end{array}$$

36
$$\begin{array}{rcrcrcrcrcr} x_1 &-& x_2 &+& 2x_3 &+& x_4 &+& x_5 &=& 9\\ x_1 &-& x_2 &-& x_3 &+& x_4 &-& x_5 &=& -1\\ -x_1 &+& x_2 &+& x_3 &-& 2x_4 &+& 2x_5 &=& 2\\ 2x_1 &-& 2x_2 &+& 3x_3 &-& x_4 &+& 2x_5 &=& 13\\ x_1 &+& 2x_2 &-& x_3 &-& 2x_4 &+& x_5 &=& -3 \end{array}$$

37
$$\begin{array}{rcrcrcrcrcr} x_1 &-& x_2 &+& x_3 &+& x_4 &+& 2x_5 &=& -2\ .\\ 2x_1 &-& x_2 &+& 2x_3 &+& 2x_4 &+& x_5 &=& -2\\ 2x_1 &+& x_2 &-& 2x_3 &+& x_4 &+& 2x_5 &=& 4\\ -x_1 &+& x_2 &-& x_3 &+& 2x_4 &+& x_5 &=& 2\\ -2x_1 &+& x_2 &-& x_3 && &+& 2x_5 &=& 0 \end{array}$$

38
$$\begin{array}{rcrcrcrcrcr} x_1 &-& x_2 &+& 2x_3 &+& 3x_4 &+& x_5 &=& 1\\ -x_1 &+& 2x_2 &+& x_3 && &+& 3x_5 &=& 5\\ -2x_1 &+& x_2 &+& x_3 &-& x_4 &+& 2x_5 &=& 0\\ 2x_1 &-& 3x_2 &+& 3x_3 &-& x_4 &+& x_5 &=& -7\\ 3x_1 &-& x_2 &+& 2x_3 && &+& x_5 &=& 0 \end{array}$$

39 Adecuar el programa `sistemas_gauss` de la figura 8-3, página 826, para construir una función en MATLAB, o en algún otro lenguaje de programación, que calcule determinantes de matrices cuadradas.

En los ejercicios 40 a 45 hallar la inversa, si existe, de la matriz A por medio del algoritmo 8.2 (pág. 828).

40 $A = \begin{bmatrix} 1 & -1 & 2\\ 2 & -1 & 1\\ 1 & -1 & 1 \end{bmatrix}.$

41 $A = \begin{bmatrix} 2 & -1 & 1\\ 1 & 2 & 1\\ 5 & 5 & 4 \end{bmatrix}.$

42 $A = \begin{bmatrix} 1 & -1 & 2 & 3\\ 2 & -1 & 1 & 4\\ 1 & -1 & 1 & 2\\ 1 & -1 & 4 & 5 \end{bmatrix}.$

43 $A = \begin{bmatrix} 1 & -1 & 2 & 3\\ 2 & -1 & 1 & 4\\ 1 & -1 & 1 & 2\\ 3 & -3 & 6 & 10 \end{bmatrix}.$

44 $A = \begin{bmatrix} 1 & -1 & 1 & 2 & 1\\ -1 & 2 & 1 & 1 & 2\\ -2 & 1 & 1 & 1 & 2\\ 1 & 1 & -1 & 2 & -3\\ 3 & -1 & 1 & 2 & 1 \end{bmatrix}.$

45 $A = \begin{bmatrix} 1 & -1 & 2 & 3 & 2\\ -1 & 2 & 2 & 1 & 4\\ 2 & -2 & 5 & 3 & 1\\ 1 & -1 & 2 & 2 & 1\\ 3 & -3 & 6 & 9 & 7 \end{bmatrix}.$

En los ejercicios 46 a 53 hallar la factorización LU de la matriz A de acuerdo a como se explicó en el apartado 8.2.3 (cfr. ejemplo 8.9, página 834).

46 $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 2 \\ 3 & 1 & 2 \end{bmatrix}.$

47 $A = \begin{bmatrix} 1 & 2 & -3 \\ 1 & 1 & 1 \\ 3 & 8 & -7 \end{bmatrix}.$

48 $A = \begin{bmatrix} 3 & -1 & 4 \\ 2 & 1 & 3 \\ -2 & 1 & 2 \end{bmatrix}.$

49 $A = \begin{bmatrix} 1 & 2 & -1 & 5 \\ -1 & -1 & 2 & 1 \\ 3 & 4 & -2 & 5 \\ 1 & 1 & 3 & 2 \end{bmatrix}.$

50 $A = \begin{bmatrix} 1 & 2 & -4 & 7 \\ -1 & -1 & 2 & -3 \\ 3 & 4 & 2 & 5 \\ 1 & 1 & 3 & -6 \end{bmatrix}.$

51 $A = \begin{bmatrix} 2 & -1 & 3 & 1 \\ 1 & -1 & 2 & 1 \\ -1 & 3 & 2 & 1 \\ -1 & 2 & 3 & 1 \end{bmatrix}.$

52 $A = \begin{bmatrix} 1 & -1 & 3 & 1 & 2 \\ 1 & -2 & -3 & 2 & -1 \\ 3 & 2 & 8 & 4 & -2 \\ 2 & -4 & 4 & 1 & 3 \\ -2 & 3 & 2 & 1 & 4 \end{bmatrix}.$

53 $A = \begin{bmatrix} 1 & 2 & 1 & 1 & 4 \\ 1 & 1 & 1 & -1 & 6 \\ 2 & -2 & 1 & -2 & 4 \\ 2 & -1 & 2 & 1 & 4 \\ 4 & -2 & 2 & 2 & 1 \end{bmatrix}.$

En los ejercicios 54 a 57 encontrar una matriz de permutación P, una matriz triangular inferior L y una matriz triangular superior U tal que $PA = LU$. (cfr. ejemplo 8.11, pág. 837 y la discusión que le antecede). La respuesta puede no ser única.

54 $A = \begin{bmatrix} 1 & -1 & 2 & -1 \\ 2 & -2 & 3 & -1 \\ 1 & -1 & 1 & 2 \\ 3 & -2 & 5 & -2 \end{bmatrix}.$

55 $A = \begin{bmatrix} 1 & 3 & -1 & 2 \\ 2 & 6 & -1 & 3 \\ 1 & 3 & -2 & 4 \\ 3 & 7 & -2 & 4 \end{bmatrix}.$

56 $A = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & -2 & 1 & 3 \\ 1 & 2 & -2 & 4 \\ 4 & -4 & 3 & 5 \end{bmatrix}.$

57 $A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ -1 & -2 & 1 & 2 \\ 2 & 3 & 6 & 3 \\ 1 & 1 & 2 & 1 \end{bmatrix}.$

Utilizar las factorizaciones LU de las matrices de los ejercicios 46, 51, 54 y 57, para resolver los sistemas lineales con la misma matriz de coeficientes contenidos en los ejercicios 58 a 61 (cfr. ejemplo 8.10 y la discusión ulterior al ejemplo 8.11).

58
$$\begin{aligned} x_1 + x_2 + 2x_3 &= 1, 2, 5 \\ 2x_1 + 3x_2 + 2x_3 &= 2, 1, 7 \\ 3x_1 + x_2 + 2x_3 &= -5, 4, 5 \end{aligned}$$

59
$$\begin{aligned} 2x_1 - x_2 + 3x_3 + x_4 &= 3, 5, 8 \\ x_1 - x_2 + 2x_3 + x_4 &= 2, 4, 6 \\ -x_1 + 3x_2 + 2x_3 + x_4 &= 2, -2, 0 \\ -x_1 + 2x_2 + 3x_3 + x_4 &= 0, -1, 2 \end{aligned}$$

60
$$\begin{array}{rcrcrcrcl} x_1 & - & x_2 & + & 2x_3 & - & x_4 & = & -1, 5, 10 \\ 2x_1 & - & 2x_2 & + & 3x_3 & - & x_4 & = & -1, 11, 18 \\ x_1 & - & x_2 & + & x_3 & + & 2x_4 & = & 0, 10, 12 \\ 3x_1 & - & 2x_2 & + & 5x_3 & - & 2x_4 & = & -2, 13, 26 \end{array}$$

61
$$\begin{array}{rcrcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & + & x_4 & = & -6, -17, -4 \\ -x_1 & - & 2x_2 & + & x_3 & + & 2x_4 & = & -6, -2, 5 \\ 2x_1 & + & 3x_2 & + & 6x_3 & + & 3x_4 & = & -14, -34, -6 \\ x_1 & + & x_2 & + & 2x_3 & + & x_4 & = & -5, -12, -3 \end{array}$$

Resolver los sistemas de los ejercicios 62 a 65 *(a)* con el método de Gauss (algoritmo 8.1) y *(b)* con la estrategia de pivote parcial mediante aritmética de punto flotante con redondeo a tres cifras significativas.

62
$$\begin{array}{rcrcl} 0.05x_1 & - & 15.2x_2 & = & 15.7 \\ 6.35x_1 & + & 7.4x_2 & = & 56.1 \end{array}$$

64
$$\begin{array}{rcrcl} 0.04x_1 & + & 57.3x_2 & = & 57.7 \\ 8.32x_1 & - & 6.3x_2 & = & 76.9 \end{array}$$

63
$$\begin{array}{rcrcl} -15.2x_1 & + & 0.05x_2 & = & 15.7 \\ 7.4x_1 & + & 6.35x_2 & = & 56.1 \end{array}$$

65
$$\begin{array}{rcrcl} 57.3x_1 & + & 0.04x_2 & = & 57.7 \\ -6.3x_1 & + & 8.32x_2 & = & 76.9 \end{array}$$

En los ejercicios 66 a 71 resolver los sistemas lineales $A\vec{x} = \vec{b}$ por el método de Gauss mediante la estrategia de pivote parcial.

66
$$\begin{bmatrix} -2 & 1 & 3 \\ 3 & -2 & 3 \\ 5 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 16 \\ 5 \\ 5 \end{bmatrix}.$$

67
$$\begin{bmatrix} 1 & 2 & -3 \\ -7 & -1 & 2 \\ 10 & 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -15 \\ 33 \end{bmatrix}.$$

68
$$\begin{bmatrix} 1 & 6 & 3 \\ -3 & 3 & 1 \\ 9 & 12 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5 \\ -6 \\ -3 \end{bmatrix}.$$

69
$$\begin{bmatrix} 1 & 2 & 3 & 1 \\ -2 & -1 & -2 & 1 \\ 7 & 1 & 2 & -1 \\ 3 & 5 & 4 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ -7 \\ 12 \\ -2 \end{bmatrix}.$$

70
$$\begin{bmatrix} 2 & -2 & 2 & 5 \\ 3 & 2 & -1 & 3 \\ 1 & 1 & 3 & 2 \\ 9 & 3 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 9 \\ 4 \\ 2 \\ 5 \end{bmatrix}.$$

71
$$\begin{bmatrix} -2 & -1 & -5 & -1 \\ 3 & 2 & 4 & 1 \\ 1 & 3 & 3 & 2 \\ 6 & 5 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -5 \\ 4 \\ 7 \\ 2 \end{bmatrix}.$$

72 Resolver el sistema lineal

$$\begin{bmatrix} 2.13 & -3.22 & 1.02 \\ 4.13 & 10.3 & -1.12 \\ 1.07 & 0.885 & 0.832 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1.11 \\ -5.25 \\ -0.238 \end{bmatrix}$$

por el método de Gauss con la estrategia de pivote parcial escalado con aritmética de punto flotante redondeando a tres cifras significativas.

73 Resolver el sistema lineal

$$\begin{bmatrix} 3.12 & -5.23 & 0.821 \\ 4.13 & 11.3 & -1.13 \\ 1.09 & 0.896 & 0.833 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8.35 \\ -7.17 \\ 0.194 \end{bmatrix}$$

por el método de Gauss con la estrategia de pivote parcial escalado con aritmética de punto flotante redondeando a tres cifras significativas.

En los ejercicios 74 a 79 resolver los sistemas lineales por el método de Gauss con la estrategia de pivote parcial escalado.

74 $$\begin{bmatrix} -2 & 3 & 3 \\ 5 & -2 & 1 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ -6 \end{bmatrix}.$$

75 $$\begin{bmatrix} 3 & -1 & 4 \\ 2 & -7 & 9 \\ -3 & 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -8 \\ -18 \\ -20 \end{bmatrix}.$$

76 $$\begin{bmatrix} 4 & -2 & 3 & 5 \\ 5 & -6 & 7 & 2 \\ 2 & -3 & 1 & 9 \\ -3 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -21 \\ -24 \\ -20 \\ -8 \end{bmatrix}.$$

77 $$\begin{bmatrix} 2 & -2 & -3 & 7 \\ 9 & -6 & 3 & 1 \\ 2 & 3 & -1 & 5 \\ -4 & -3 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -15 \\ 2 \\ -18 \\ -3 \end{bmatrix}.$$

78 $$\begin{bmatrix} -3 & 1 & -2 & 2 \\ 9 & -2 & 1 & -1 \\ -2 & -3 & 1 & -9 \\ 5 & 2 & 3 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 9 \\ 6 \\ -35 \\ 24 \end{bmatrix}.$$

79 $$\begin{bmatrix} 5 & 2 & -3 & 2 \\ 4 & 1 & -3 & 5 \\ 4 & -2 & 4 & 5 \\ 3 & 1 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -8 \\ -20 \\ -40 \\ -20 \end{bmatrix}.$$

80 Resolver el sistema lineal

$$\begin{bmatrix} 21.33 & 18.21 & -30.27 \\ -2.23 & 32.00 & -1.11 \\ -1.15 & -11.02 & -2.38 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -51.27 \\ 26.44 \\ -19.31 \end{bmatrix}$$

por el método de Gauss con la estrategia de pivote total y con aritmética de punto flotante con redondeo a cuatro cifras significativas.

81 Resolver el sistema lineal

$$\begin{bmatrix} 18.24 & 21.22 & -40.38 \\ -3.11 & 31.00 & -1.24 \\ -1.17 & 10.02 & -2.27 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -60.46 \\ 55.17 \\ 12.06 \end{bmatrix}$$

por el método de Gauss con la estrategia de pivote total y con aritmética de punto flotante con redondeo a cuatro cifras significativas.

En los ejercicios 82 a 87 resolver los sistemas lineales por medio del método de Gauss con la estrategia de pivote total.

82 $$\begin{bmatrix} 2 & -24 & 3 \\ 2 & 1 & -1 \\ 3 & 11 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 31 \\ 1 \\ 25 \end{bmatrix}.$$

83 $$\begin{bmatrix} 1 & -48 & -3 \\ -12 & -1 & -4 \\ -3 & 12 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} -78 \\ -18 \\ 20 \end{bmatrix}.$$

84 $$\begin{bmatrix} 7 & 2 & 1 \\ -1 & 1 & 9 \\ 2 & 3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} -21 \\ 33 \\ -11 \end{bmatrix}.$$

85 $$\begin{bmatrix} 2 & -2 & 1 & 1 \\ 1 & -1 & -45 & -1 \\ -15 & 12 & 2 & -4 \\ 2 & -1 & 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 40 \\ -12 \\ 9 \end{bmatrix}.$$

86 $$\begin{bmatrix} 2 & -5 & 30 & -10 \\ 1 & -3 & 9 & -2 \\ 4 & 3 & -3 & 50 \\ 11 & -10 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2 \\ -5 \\ -48 \\ -35 \end{bmatrix}.$$

87 $$\begin{bmatrix} 1 & -2 & 2 & 55 & 2 \\ 15 & -23 & 21 & 1 & -3 \\ 2 & -1 & 3 & 18 & -9 \\ -2 & 3 & -1 & 4 & -10 \\ 1 & -1 & 5 & 8 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -46 \\ 7 \\ -47 \\ -45 \\ 26 \end{bmatrix}.$$

Métodos iterativos (respuestas en páginas 1100-1101)

Sean **E** un espacio vectorial y S un subespacio de **E**. Se dice que $p \in S$ es un punto fijo de una función $g : S \to \mathbf{E}$ si $g(p) = p$. En los ejercicios 88 a 95 hallar los puntos fijos de la función g.

88 $g(x) = x^3$, $\mathbf{E} = S = \mathbb{R}$.

89 $g(x) = x^2 - 6$, $\mathbf{E} = S = \mathbb{R}$.

90 $g(x) = x^3 - x + 1$, $\mathbf{E} = S = \mathbb{R}$.

91 $g(x) = \operatorname{sen} x$. $\mathbf{E} = S = \mathbb{R}$.

92 $g(x,y) = (2x - y, x - y)$, $\mathbf{E} = S = \mathbb{R}^2$.

93 $g(x,y) = (2x + y, x + 2y)$, $\mathbf{E} = S = \mathbb{R}$.

94 $g(f) = f'$, $\mathbf{E} = C(-\infty,\infty)$, $S = C^1(-\infty,\infty)$.

95 $g(A) = A^2$, $\mathbf{E} = S = \mathfrak{M}_2(\mathbb{R})$.

96 Sea f una función con valores reales definida en un intervalo J. Se dice que f es una función de Lipschitz en J si existe una constante $L > 0$ tal que

$$|f(x) - f(y)| \le L|x - y|$$

para todo $x, y \in J$ (cfr. nota 8.3, pág. 852). En tal caso L es una constante de Lipschitz para la función f en el intervalo J. Demostrar que si la función f tiene primera derivada acotada en el intervalo J; esto es, existe $M > 0$ tal que $|f'(x)| \le M$ para todo $x \in J$, entonces f es una función de Lipschitz en J y M es una constante de Lipschitz para f en ese intervalo.

97 En relación con el ejercicio 96, demostrar que la función $f(x) = \operatorname{sen}(x)$ es una función de Lipschitz en $J = \mathbb{R}$. Determinar la menor constante de Lipschitz para esta función en J.

98 En relación con el ejercicio 96, demostrar que la función $f(x) = \arctan(x)$ es una función de Lipschitz en $J = \mathbb{R}$. Determinar la menor constante de Lipschitz para esta función en J.

99 La función $f(x) = 2^{-x}$ tiene un único punto fijo p en el intervalo $J = [0,1]$.

(a) Mostrar que $|f'(x)| \le \frac{\ln 2}{2}$ para todo $x \in J$.

(b) Sea $\alpha = \frac{\ln 2}{2}$, utilizar la información del inciso anterior y la fórmula (8.17) del corolario 8.1 (cfr. pág. 852) para hallar la primera iteración de punto fijo que aproxima a p con un error menor a 10^{-3} si $\vec{x}_0 = 0.5$.

100 Sean $\|\cdot\|^{(\infty)}$ y $\|\cdot\|^{(1)}$ las normas matriciales naturales subordinadas a las normas vectoriales $\|\cdot\|_\infty$ y $\|\cdot\|^1$ de $\mathbb{R}^n$, respectivamente. Mostrar que para toda $A \in \mathfrak{M}_n(\mathbb{R})$ se tiene:

(a) $\|A^t\|^{(1)} = \|A\|^{(\infty)}$.

(b) $\|A^t\|^{(\infty)} = \|A\|^{(1)}$.

101 Sea $A \in \mathfrak{M}_n(\mathbb{R})$, entonces la matriz A^tA es simétrica y semidefinida positiva; esto es, $\vec{x}^t(A^tA)\vec{x} \geq 0$ para todo $\vec{x} \in \mathbb{R}^n$. Sea $\|\cdot\|_2$ la norma vectorial euclidiana; es decir, $\|\vec{x}\| = \sqrt{\vec{x}\cdot\vec{x}}$, demostrar que la norma matricial natural subordinada a ella, $\|\cdot\|^{(2)}$, está dada por

$$\|A\|^{(2)} = (\sigma_{\text{máx}}(A^tA))^{1/2}$$

donde

$$\sigma_{\text{máx}}(A^tA) = \text{máx}\{\sigma : \sigma \text{ es valor propio de } A^tA\}.$$

102 Sea $\|\cdot\|^{(2)}$ la norma matricial natural que proviene de la norma vectorial euclidiana $\|\cdot\|_2$, probar que

$$\|A\|^{(2)} = \max_{\|\vec{y}\|_2=\|\vec{x}\|_2=1} \vec{y}^tA\vec{x}.$$

103 Mostrar que $\|A\|^{(2)} = \|A^t\|^{(2)}$.

104 Se define, para cada $A \in \mathfrak{M}_n(\mathbb{R})$,

$$\|A\|_F = \sqrt{\text{tra}(A^tA)}.$$

Mostrar que $\|\cdot\|_F$ es una norma en el espacio $\mathfrak{M}_n(\mathbb{R})$. $\|\cdot\|_F$ se llama norma de Frobenius.

105 Sea $\|\cdot\|_F$ la norma de Frobenius definida en el ejercicio anterior. Probar que, para cada $A \in \mathfrak{M}_n(\mathbb{R})$,

$$\|A\|_F = \left(\sum_{i=1}^{n} \sigma_i^2\right)^{1/2}$$

donde σ_i, $i = 1, 2, \ldots, n$, son los valores propios de la matriz simétrica A^tA.

106 Sea $\|\cdot\|_F$ la norma de Frobenius de los dos ejercicios previos, mostrar que:

(a) $\|A\vec{x}\|_2 \leq \|A\|_F\,\|\vec{x}\|_2$ para todo $\vec{x} \in \mathbb{R}^n$ y para toda $A \in \mathfrak{M}_n(\mathbb{R})$.

(b) $\|AB\|_F \leq \|A\|_F\,\|B\|_F$ para todo par de matrices $A, B \in \mathfrak{M}_n(\mathbb{R})$.

107 Calcular $\|I_n\|_F$.

En los ejercicios 108 a 115 encontrar la primera iteración $\vec{x}_k$ que, mediante el método de Jacobi con cota para el error (algoritmo 8.5, pág. 866), aproxima a la única solución del sistema lineal con un error inferior al valor ε, tomando como vector inicial del ciclo iterativo al vector $\vec{x}_0$ (cfr. ejemplo 8.22 pág. 862). Hacer los cálculos con calculadora sin utilizar programas de computación.

108

$$\begin{array}{rrrrrcr}
50x_1 & & +\ 2x_3 & & -\ 4x_5 & = & -48 \\
 & 40x_2 & & +\ 3x_4 & & = & 37 \\
 & -\ 2x_2 & +\ 60x_3 & & -\ x_5 & = & -61 \\
2x_1 & & & +\ 70x_4 & & = & -72 \\
 & 4x_2 & & -2x_4 & +\ 80x_5 & = & -74
\end{array}$$

$\varepsilon = 0.01$, $\vec{x}_0 = (-2, 0, -2, 0, 0)$.

109
$$\begin{array}{rcrcrcrcrcr} -75x_1 & & & - & 3x_3 & + & x_4 & - & x_5 & = & -73 \\ -2x_1 & - & 65x_2 & - & x_3 & & & + & 2x_5 & = & -133 \\ -x_1 & + & x_2 & - & 85x_3 & + & 3x_4 & & & = & 80 \\ -x_1 & - & 2x_2 & - & 2x_3 & + & 80x_4 & & & = & -163 \\ -3x_1 & + & 2x_2 & & & & -2x_4 & + & 75x_5 & = & -70 \end{array}$$

$\varepsilon = 0.01$, $\vec{x}_0 = (0, 1, 0, 1, 0)$.

110
$$\begin{array}{rcrcrcrcrcr} 95x_1 & - & 4x_2 & & & - & 3x_4 & + & x_5 & = & -100 \\ -2x_1 & - & 75x_2 & & & - & 2x_4 & + & x_5 & = & -220 \\ & - & 3x_2 & + & 80x_3 & + & 4x_4 & - & 2x_5 & = & -175 \\ x_1 & + & 5x_2 & - & 3x_3 & - & 85x_4 & & & = & 360 \\ x_1 & - & 5x_2 & & & + & 3x_4 & + & 70x_5 & = & -378 \end{array}$$

$\varepsilon = 0.001$, $\vec{x}_0 = (0, 2, -1, -3, -4)$.

111
$$\begin{array}{rcrcrcrcrcr} 50x_1 & + & 3x_2 & - & 4x_3 & & & + & 2x_5 & = & -59 \\ & + & 79x_2 & - & x_3 & - & x_4 & + & 3x_5 & = & 228 \\ -x_1 & - & 4x_2 & + & 95x_3 & & & - & 3x_5 & = & 372 \\ 2x_1 & - & x_2 & + & 2x_3 & - & 80x_4 & & & = & -157 \\ 3x_1 & - & 3x_2 & & & - & 2x_4 & + & 75x_5 & = & -91 \end{array}$$

$\varepsilon = 0.01$, $\vec{x}_0 = (-2, 2, 3, 1, 0)$.

112
$$\begin{array}{rcrcrcrcrcr} 80x_1 & + & 3x_2 & - & x_3 & & & + & x_5 & = & 78 \\ & + & 85x_2 & - & x_3 & - & x_4 & - & 2x_5 & = & -84 \\ -x_1 & - & x_2 & + & 95x_3 & & & - & 3x_5 & = & -187 \\ x_1 & - & x_2 & + & 2x_3 & - & 90x_4 & & & = & -272 \\ -x_1 & - & 2x_2 & & & + & x_4 & - & 90x_5 & = & 94 \end{array}$$

$\varepsilon = 0.001$, $\vec{x}_0 = (0, -2, -1, 2, 0)$.

113
$$\begin{array}{rcrcrcrcrcr} 90x_1 & + & x_2 & & & + & x_4 & - & x_5 & = & 173 \\ & - & 85x_2 & + & 2x_3 & - & x_4 & - & x_5 & = & 167 \\ & - & 3x_2 & + & 100x_3 & + & 3x_4 & - & 3x_5 & = & -309 \\ 3x_1 & - & 2x_2 & - & x_3 & + & 80x_4 & & & = & -307 \\ -x_1 & - & 2x_2 & - & x_3 & & & - & 200x_5 & = & -195 \end{array}$$

$\varepsilon = 0.001$, $\vec{x}_0 = (1, -1, -2, -5, 0)$.

114
$$\begin{array}{rcrcrcrcrcrcr} 90x_1 & - & x_2 & + & 3x_3 & & & - & x_5 & + & x_6 & = & -272 \\ -3x_1 & + & 100x_2 & & & - & x_4 & - & x_5 & + & 2x_6 & = & 206 \\ & - & x_2 & + & 200x_3 & - & 5x_4 & + & 2x_5 & - & 4x_6 & = & 218 \\ x_1 & & & + & 3x_3 & - & 95x_4 & - & x_5 & + & x_6 & = & 187 \\ x_1 & - & x_2 & - & x_3 & + & 2x_4 & - & 300x_5 & & & = & -310 \\ 2x_1 & - & 3x_2 & + & 2x_3 & - & x_4 & & & - & 100x_6 & = & 192 \end{array}$$

$\varepsilon = 0.0001$, $\vec{x}_0 = (-2, 1, 0, -1, 0, -1)$.

115
$$\begin{array}{rcrcrcrcrcrcr} 100x_1 & + & 4x_2 & & & - & 5x_4 & - & x_5 & + & x_6 & = & -332 \\ -x_1 & + & 300x_2 & & & - & x_4 & - & x_5 & + & 2x_6 & = & -610 \\ x_1 & - & x_2 & - & 400x_3 & + & 3x_4 & - & 2x_5 & & & = & -387 \\ & - & 2x_2 & + & 4x_3 & - & 550x_4 & & & + & x_6 & = & -2197 \\ -2x_1 & - & 3x_2 & + & 2x_3 & - & x_4 & - & 200x_5 & & & = & 210 \\ x_1 & - & x_2 & + & x_3 & - & 5x_4 & & & - & 300x_6 & = & 1480 \end{array}$$

$\varepsilon = 0.001$, $\vec{x}_0 = (-2, -1, 2, 3, -2, -6)$.

En los ejercicios 116 a 121 calcular la primera aproximación $\vec{x}_k$, de la única solución del sistema lineal $A\vec{x} = \vec{b}$, que satisface $\|\vec{x}_k - \vec{x}_{k-1}\| < \delta$, mediante el método de Jacobi con margen de tolerancia entre las diferencias de iteraciones sucesivas (cfr. algoritmo 8.6, pág. 872 y el ejemplo 8.27), tomando como aproximación inicial el vector $\vec{x}_0$. Calcular también $\|T\|^{\infty}$, donde T es la matriz definida para cada sistema por la relación (8.31) de la página 866.

116 $A = \begin{bmatrix} 20 & 0 & 1 \\ 0 & 50 & -1 \\ 1 & 1 & 80 \end{bmatrix}, \vec{b} = \begin{bmatrix} -19 \\ 99 \\ 81 \end{bmatrix}, \vec{x}_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \delta = 0.001.$

117 $A = \begin{bmatrix} 10 & -1 & 0 \\ 0 & 20 & -1 \\ 1 & -1 & 30 \end{bmatrix}, \vec{b} = \begin{bmatrix} 12 \\ -37 \\ -87 \end{bmatrix}, \vec{x}_0 = \begin{bmatrix} 0 \\ -1 \\ -2 \end{bmatrix}, \delta = 0.01.$

118 $A = \begin{bmatrix} 4 & -3 & 0 & 1 & 0 \\ 5 & 10 & -2 & 0 & 0 \\ 0 & -2 & 3 & 1 & 0 \\ -1 & 0 & 4 & 6 & 0 \\ 2 & 1 & -1 & 0 & 5 \end{bmatrix}, \vec{b} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ -3 \\ -4 \end{bmatrix}, \vec{x}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \delta = 0.01.$

119 $A = \begin{bmatrix} -5 & 2 & -1 & 0 & 1 \\ 2 & 15 & -1 & 0 & 1 \\ -1 & 0 & 8 & 3 & 0 \\ -1 & 0 & 0 & 9 & 1 \\ 0 & -1 & 2 & 4 & 10 \end{bmatrix}, \vec{b} = \begin{bmatrix} 7 \\ 26 \\ 24 \\ -24 \\ 14 \end{bmatrix}, \vec{x}_0 = \begin{bmatrix} 0 \\ 1 \\ 3 \\ -2 \\ 1 \end{bmatrix}, \delta = 0.01.$

120 $A = \begin{bmatrix} 10 & 2 & -1 & 0 & -1 \\ 2 & 20 & -2 & 0 & 2 \\ 1 & -2 & 30 & 0 & -2 \\ 1 & 0 & -4 & 5 & 20 \\ 1 & 0 & -1 & 1 & 40 \end{bmatrix}, \vec{b} = \begin{bmatrix} 12 \\ 2 \\ -27 \\ -10 \\ -37 \end{bmatrix}, \vec{x}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \delta = 0.01.$

121 $A = \begin{bmatrix} 10 & 2 & -1 & 0 & -1 \\ 2 & 20 & -2 & 0 & 2 \\ 1 & -2 & 30 & 0 & -2 \\ 1 & 0 & -4 & 20 & 5 \\ 1 & 0 & -1 & 1 & 40 \end{bmatrix}, \vec{b} = \begin{bmatrix} -7 \\ 28 \\ 89 \\ -63 \\ -86 \end{bmatrix}, \vec{x}_0 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ -1 \\ -1 \end{bmatrix}, \delta = 0.01.$

En los ejercicios 122 a 127 encontrar la primera iteración $\vec{x}_k$ que, mediante el método de Richardson con cota para el error (algoritmo 8.7, pág. 883), aproxima a la única solución del sistema lineal $A\vec{x} = \vec{b}$ con un error inferior al valor ε, tomando como vector inicial del ciclo iterativo al vector $\vec{x}_0$ (cfr. ejemplo 8.33, pág. 882). Hacer los cálculos con calculadora sin utilizar programas de computación.

122 $A = \begin{bmatrix} 1 & 0 & 1/40 \\ -1/40 & 1 & 0 \\ -1/80 & 0 & 1 \end{bmatrix}, \vec{b} = \begin{bmatrix} 78 \\ 158 \\ -81 \end{bmatrix}, \vec{x}_0 = \begin{bmatrix} 70 \\ 150 \\ -60 \end{bmatrix}, \varepsilon = 0.01.$

123 $A = \begin{bmatrix} 1 & -1/100 & 1/50 \\ 0 & 1 & -1/50 \\ 1/100 & 0 & 1 \end{bmatrix}, \vec{b} = \begin{bmatrix} -97 \\ 96 \\ 199 \end{bmatrix}, \vec{x}_0 = \begin{bmatrix} 80 \\ 80 \\ 150 \end{bmatrix}, \varepsilon = 0.001.$

124 $A = \begin{bmatrix} 1 & 0 & -1/40 & 1/80 \\ 0 & 1 & -1/80 & 0 \\ 0 & -1/40 & 1 & 0 \\ 0 & -1/80 & 1/40 & 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} -152 \\ -318 \\ -152 \\ 320 \end{bmatrix}$, $\vec{x}_0 = \begin{bmatrix} -140 \\ -300 \\ -160 \\ 300 \end{bmatrix}$, $\varepsilon = 0.01$.

125 $A = \begin{bmatrix} 1 & -1/200 & 0 & 1/200 \\ 0 & 1 & 1/200 & -1/200 \\ 1/400 & 0 & 1 & -1/400 \\ 0 & -1/200 & 1/200 & 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} -788 \\ -1598 \\ 1196 \\ 814 \end{bmatrix}$, $\vec{x}_0 = \begin{bmatrix} -900 \\ -1500 \\ 100 \\ 700 \end{bmatrix}$, $\varepsilon = 0.01$.

126 $A = \begin{bmatrix} 1 & 0 & 1/80 & 1/40 \\ 1/80 & 1 & 0 & 0 \\ -1/40 & 0 & 1 & -1/80 \\ 0 & 1/40 & 1/40 & 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 83 \\ -79 \\ 77 \\ 80 \end{bmatrix}$, $\vec{x}_0 = \begin{bmatrix} 75 \\ -75 \\ 75 \\ 75 \end{bmatrix}$, $\varepsilon = 0.001$.

127 $A = \begin{bmatrix} 1 & 0 & 0 & 1/250 & -1/250 \\ 0 & 1 & 1/500 & -1/500 & 0 \\ -1/500 & 0 & 1 & 1/500 & -1/250 \\ 0 & 0 & -1/500 & 1 & 0 \\ 1/500 & 0 & 0 & -1/500 & 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 506 \\ -504 \\ -997 \\ 1002 \\ -501 \end{bmatrix}$, $\vec{x}_0 = \begin{bmatrix} 600 \\ -600 \\ -800 \\ 800 \\ -400 \end{bmatrix}$, $\varepsilon = 0.001$.

En los ejercicios 128 a 131 calcular la primera aproximación $\vec{x}_k$, de la única solución del sistema lineal $A\vec{x} = \vec{b}$, que satisface $\|\vec{x}_k - \vec{x}_{k-1}\| < \delta$, mediante el método de Richardson con margen de tolerancia entre las diferencias de iteraciones sucesivas (cfr. algoritmo 8.8, pág. 886), tomando como aproximación inicial el vector $\vec{x}_0$. Calcular también $\|T\|^{\infty}$, donde $T = I_n - A$ (cfr. apartado 8.3.4, página 880).

128 $A = \begin{bmatrix} 1 & -1/2 & 1/2 \\ 0 & 1 & 1/8 \\ 0 & -1/4 & 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 20 \\ -6 \\ 18 \end{bmatrix}$, $\vec{x}_0 = \begin{bmatrix} 7 \\ -7 \\ 15 \end{bmatrix}$, $\delta = 0.01$.

129 $A = \begin{bmatrix} 1 & -1/3 & 2/3 \\ 0 & 1 & -1/3 \\ 1/3 & 0 & 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 11 \\ -9 \\ 10 \end{bmatrix}$, $\vec{x}_0 = \begin{bmatrix} 4 \\ -7 \\ 8 \end{bmatrix}$, $\delta = 0.01$.

130 $A = \begin{bmatrix} 1 & -1/50 & 0 & 1/25 \\ 0 & 1 & -1/25 & 0 \\ 0 & -1/50 & 1 & 1/100 \\ 0 & -1/50 & 1/100 & 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 94 \\ -112 \\ 300 \\ -195 \end{bmatrix}$, $\vec{x}_0 = \begin{bmatrix} 90 \\ -90 \\ 310 \\ -210 \end{bmatrix}$, $\delta = 0.01$.

131 $A = \begin{bmatrix} 1 & -1/200 & 0 & 1/100 \\ 0 & 1 & 1/100 & 1/200 \\ 0 & -1/200 & 1 & -1/200 \\ 1/100 & 1/200 & 0 & 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} -196 \\ -405 \\ -599 \\ 196 \end{bmatrix}$, $\vec{x}_0 = \begin{bmatrix} -210 \\ -390 \\ -610 \\ 200 \end{bmatrix}$, $\delta = 0.001$.

En los ejercicios 132 a 139 calcular la primera aproximación $\vec{x}_k$, de la única solución del sistema lineal $A\vec{x} = \vec{b}$, que satisface $\|\vec{x}_k - \vec{x}_{k-1}\| < \delta$, mediante el método de Gauss-Seidel con margen de tolerancia entre las diferencias de iteraciones sucesivas (cfr. algoritmo 8.9, pág. 891), tomando como aproximación inicial el vector $\vec{x}_0$.

132 $A\vec{x} = \vec{b}$ es el sistema del ejercicio 108, $\vec{x}_0 = (0, 0, 0, 0, 0)$, $\delta = 0.001$.

133 $A\vec{x} = \vec{b}$ es el sistema del ejercicio 110, $\vec{x}_0 = (0, 2, -1, -3, -4)$, $\delta = 0.01$.

134 $A\vec{x} = \vec{b}$ es el sistema del ejercicio 109, $\vec{x}_0 = (0, 1, 0, 1, 0)$, $\delta = 0.001$.

135 $A\vec{x} = \vec{b}$ es el sistema del ejercicio 112, $\vec{x}_0 = (0, -2, -1, -2, 0)$, $\delta = 0.01$.

136 $A\vec{x} = \vec{b}$ es el sistema del ejercicio 111, $\vec{x}_0 = (-2, 2, 3, 1, 0)$, $\delta = 0.001$.

137 $A\vec{x} = \vec{b}$ es el sistema del ejercicio 114, $\vec{x}_0 = (-2, 1, 0, -1, 0, -1)$, $\delta = 0.01$.

138 $A\vec{x} = \vec{b}$ es el sistema del ejercicio 115, $\vec{x}_0 = (-2, -1, 2, 3, -2, -6)$, $\delta = 0.001$.

139 $A\vec{x} = \vec{b}$ es el sistema del ejercicio 113, $\vec{x}_0 = (1, -1, -2, -5, 0)$, $\delta = 0.0001$.

140 Repetir el ejercicio 116 pero con el método de Gauss-Seidel y sin calcular $\|T\|^{\infty}$.

141 Repetir el ejercicio 117 pero con el método de Gauss-Seidel y sin calcular $\|T\|^{\infty}$.

142 Repetir el ejercicio 118 pero con el método de Gauss-Seidel y sin calcular $\|T\|^{\infty}$.

143 Repetir el ejercicio 119 pero con el método de Gauss-Seidel y sin calcular $\|T\|^{\infty}$.

144 Repetir el ejercicio 120 pero con el método de Gauss-Seidel y sin calcular $\|T\|^{\infty}$.

145 Repetir el ejercicio 121 pero con el método de Gauss-Seidel y sin calcular $\|T\|^{\infty}$.

En los ejercicios 146 a 151 aplique el algorimo 8.10, página 898, para hallar la inversa de la matriz B con un margen de tolerancia δ.

146 $B = \begin{bmatrix} 1.0000 & 0.1000 \\ 0.2000 & 1.0000 \end{bmatrix}$, $\delta = 0.01$.

147 $B = \begin{bmatrix} 0.8800 & 0.0100 \\ -0.0100 & 0.8900 \end{bmatrix}$, $\delta = 0.01$.

148 $B = \begin{bmatrix} 1.0000 & 0.0100 & -0.0200 \\ 0.0000 & 1.0000 & -0.0200 \\ 0.0100 & -0.0100 & 1.0000 \end{bmatrix}$, $\delta = 0.001$.

149 $B = \begin{bmatrix} 0.8800 & -0.0100 & 0.0200 \\ 0.0200 & 0.9800 & -0.0200 \\ 0.0200 & -0.0100 & 0.8700 \end{bmatrix}$, $\delta = 0.01$.

150 $B = \begin{bmatrix} -0.0100 & 0.0100 & -0.0200 \\ -0.0400 & 1.0000 & -0.0300 \\ 0.0200 & -0.0300 & 1.0000 \end{bmatrix}$, $\delta = 0.001$.

151 $B = \begin{bmatrix} 0.9000 & -0.0100 & -0.0200 \\ 0.0100 & 0.9500 & -0.0300 \\ 0.0200 & -0.0100 & 0.9000 \end{bmatrix}$, $\delta = 0.001$.

Transformaciones Householder (respuestas en páginas 1101-1103)

En los ejercicios 152 a 161 utilizar las relaciones (8.56), pág. 905, para encontrar la matriz de Householder H tal, que la función $\vec{x} \mapsto H\vec{x}$ tranforma el vector $\vec{u} = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$ en un vector de la forma $(\alpha, 0, \ldots, 0) \in \mathbb{R}^n$, con $\alpha = -\text{sign}(\vec{u}_1)\,\|\vec{u}\|$. Realizar los cálculos sin usar programas de computación.

152 $\vec{u} = (1, 2, 2)$.

153 $\vec{u} = (1, -1, -2)$.

154 $\vec{u} = (4, -7, -4)$.

155 $\vec{u} = (-2, -3, 6)$.

156 $\vec{u} = (4, -2, 4)$.

157 $\vec{u} = (-8, 4, -1)$.

158 $\vec{u} = (-3, 2, 6)$.

159 $\vec{u} = (1, -1, 5, -3)$.

160 $\vec{u} = (2, -4, 5, 2)$.

161 $\vec{u} = (-4, 1, 4, 4)$.

162 Si H es una matriz de Householder de orden 2, demostrar que H tiene la forma

$$H = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

donde $a, b \in \mathbb{R}$ y $a^2 + b^2 = 1$.

En los ejercicios 163 a 170 hallar la matriz de Householder H tal que la aplicación $\vec{x} \mapsto H\vec{x}$ transforma al vector dado $\vec{u}$ en el vector $H\vec{u}$ que satisface la condición indicada. (cfr. ejemplo 8.43 y discusión que le precede, pág. 907). Realizar los cálculos sin el uso de programa alguno de computación.

163 $\vec{u} = (-8, 4, -3)$, $H\vec{u}$ tiene la tercera coordenada igual a cero.

164 $\vec{u} = (1, -1, 4, -3)$, $H\vec{u}$ tiene la última coordenada igual a cero.

165 $\vec{u} = (3, -3, 2, -6)$, $H\vec{u}$ tiene las dos últimas coordenadas iguales a cero.

166 $\vec{u} = (1, 2, -3, 6)$, $H\vec{u}$ tiene las dos últimas coordenadas iguales a cero.

167 $\vec{u} = (2, -7, -4, 4)$, $H\vec{u}$ tiene las dos últimas coordenadas iguales a cero.

168 $\vec{u} = (-4, 4, -2, 4)$, $H\vec{u}$ tiene las dos últimas coordenadas iguales a cero.

169 $\vec{u} = (1, -4, -3, -1, 3, 1)$, $H\vec{u}$ tiene las cuatro últimas coordenadas iguales a cero.

170 $\vec{u} = (1, -4, 1, -8, 4)$, $H\vec{u}$ tiene las dos últimas coordenadas iguales a cero.

171 Demostrar que si $\vec{u}$ y $\vec{x}$ son un par de vectores en $\mathbb{R}^n$, entonces

$$(\vec{u}\vec{u}^t)\vec{x} = (\vec{u}^t\vec{x})\vec{u}.$$

Sean $\vec{u} \in \mathbb{R}^n$ un vector tal que $\vec{u} \cdot \vec{u} = 1$ y $H = I_n - 2\vec{u}\vec{u}^t$ la matriz de Householder que este vector define. Sea

$$S = \text{gn}(\vec{u}).$$

Sabemos (cfr. ejercicios propuestos 95 a 101 del capítulo 4, pág. 389) que entonces $\mathbb{R}^n = S \oplus S^\perp$, donde $S^\perp$ es el subespacio ortogonal de S. Luego, todo vector $\vec{x} \in \mathbb{R}^n$ se puede escribir, de manera única, en la forma

$$\vec{x} = \vec{x}_S + \vec{x}_{S^\perp}$$

donde $\vec{x}_S \in S$ y $\vec{x}_{S^\perp} \in S^\perp$. Utilizar esta información para demostrar las afirmaciones contenidas en los ejercicios 172 a 179 relativas a H.

172 Si $\vec{x} \in S$, entonces $H\vec{x} = -\vec{x}$.

173 Si $\vec{x} \in S^\perp$, entonces $H\vec{x} = \vec{x}$.

174 Si $\vec{x} = \vec{x}_S + \vec{x}_{S^\perp}$, $\vec{x}_S \in S$ y $\vec{x}_{S^\perp} \in S^\perp$, entonces

$$H\vec{x} = -\vec{x}_S + \vec{x}_{S^\perp}$$

175 H^2 es una involución; es decir, $H^2 = I$.

176 H sólo tiene los valores propios $\lambda_1 = 1$ y $\lambda_2 = -1$.

177 $E_{\lambda_2} = S$ y $E_{\lambda_1} = S^\perp$; es decir, los espacios propios de $\lambda_2 = -1$ y $\lambda_1 = 1$ son, respectivamente, S y su complemento ortogonal.

178 $\det(H) = -1$.

179 H es una isometría relativa a la norma euclidiana $\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}}$; esto es,

$$\|H\vec{x}\| = \|\vec{x}\|\,.$$

180 Sean $\vec{u}$ y $\vec{v}$ un par de vectores en $\mathbb{R}^n$ que son distintos entre sí y unitarios; es decir, $\vec{u}\cdot\vec{u} = 1$ y $\vec{v}\cdot\vec{v} = 1$. Mostrar que existe una transformación de Householder $\vec{x} \mapsto H\vec{x}$, tal que $H\vec{u} = \vec{v}$. (Sugerencia: tratar con $H = I_n - 2\vec{w}w^t$, donde $\vec{w} = (\vec{u} - \vec{v})/\|\vec{u} - \vec{v}\|$, con $\|\cdot\|$ obviamente la norma euclidiana).

181 Si

$$A = \begin{bmatrix} 2 & -1 & -1 & -2 \\ -4 & 1 & 2 & 1 \\ 5 & 1 & -1 & 1 \\ -2 & -1 & 2 & 1 \end{bmatrix}$$

(a) Encontrar el vector unitario $\vec{u}$ tal que la transformación de Householder $\vec{x} \mapsto H\vec{x}$ transforma la primera columna $\vec{a}_1$ de la matriz A en el vector $H\vec{a}_1 = (\alpha, 0, 0, 0)$, donde $\alpha = -\|\vec{a}_1\|$.

(b) Sin determinar explícitamente la matriz H del inciso precedente, calcular el producto HA (cfr. discusión ulterior a la proposición 8.8, pág. 902).

182 Si

$$A = \begin{bmatrix} 1 & -2 & -3 & -1 \\ -3 & 1 & 2 & 1 \\ 1 & 1 & -1 & -1 \\ -5 & 2 & 2 & 1 \end{bmatrix}$$

(a) Encontrar el vector unitario $\vec{u}$ tal que la transformación de Householder $\vec{x} \mapsto H\vec{x}$ transforma la primera columna $\vec{a}_1$ de la matriz A en el vector $H\vec{a} = (\alpha, 0, 0, 0)$, donde $\alpha = -\|\vec{a}_1\|$.

(b) Sin determinar explícitamente la matriz H del inciso precedente, calcular el producto HA (cfr. discusión ulterior a la proposición 8.8, pág. 902).

En los ejercicios 183 a 187 hallar, haciendo las operaciones manualmente como se hizo en el ejemplo 8.44 de la sección 8.4.2, la factorización QR de Householder para cada una de las matrices ahí dadas.

183 $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & 1 \\ -2 & 4 & 6 \end{bmatrix}$.

184 $A = \begin{bmatrix} 2 & 1 & -1 \\ -1 & -3 & 4 \\ 2 & -4 & 3 \end{bmatrix}$.

185 $A = \begin{bmatrix} 2 & 4 & -1 \\ 3 & 27 & 1 \\ -6 & 23 & 5 \end{bmatrix}$.

186 $A = \begin{bmatrix} -3 & 1 & 2 \\ 2 & -4 & -1 \\ 6 & 3 & 4 \end{bmatrix}$

187 $A = \begin{bmatrix} 1 & -9 & -1 \\ -4 & -4 & 3 \\ 8 & -2 & -1 \end{bmatrix}$.

Utilizar las factorizaciones QR calculadas en los ejercicios 183 a 187 para resolver los sistemas lineales de los ejercicios 188 a 192. Proceder como se resolvió el sistema del ejemplo 8.44.

188 $\begin{array}{rcrcrcr} x & - & y & + & z & = & 3 \\ 2x & + & 3y & + & z & = & 1 \\ -2x & + & 4y & + & 6z & = & -8 \end{array}$

189 $\begin{array}{rcrcrcr} 2x & + & y & - & z & = & 1 \\ -x & - & 3y & + & 4z & = & 1 \\ 2x & - & 4y & + & 3z & = & 7 \end{array}$

190 $\begin{array}{rcrcrcr} 2x & + & 4y & - & z & = & 5 \\ 3x & + & 27y & + & z & = & 31 \\ -6x & + & 23y & + & 5z & = & 22 \end{array}$

191 $\begin{array}{rcrcrcr} -3x & + & y & + & 2z & = & -1 \\ 2x & - & 4y & - & z & = & -7 \\ 6x & + & 3y & + & 4z & = & -12 \end{array}$

192 $\begin{array}{rcrcrcr} x & - & 9y & - & z & = & 16 \\ -4x & - & 4y & + & 3z & = & 15 \\ 8x & - & 2y & - & z & = & -5 \end{array}$

193 Demostrar la generalización del teorema 8.9 para matrices rectangulares:

Teorema: **[Factorización QR para matrices rectangulares]** Si A es una matriz real de tamaño $m \times n$ con $m \geq n$, entonces existen dos matrices, $Q \in \mathfrak{M}_m(\mathbb{R})$ —matriz ortogonal, producto de matrices de Householder— y una matriz triangular superior $\widehat{R} \in \mathfrak{M}_n(\mathbb{R})$, tal que

$$A = QR$$

donde $R = \begin{bmatrix} \widehat{R} \\ \mathscr{O}_{m-n,n} \end{bmatrix}$.

Utilizar el resultado del ejercicio precedente para hallar las factorizaciones QR de las matrices de los ejercicios 194 y 195. No utilizar programa alguno de computación. Realizar los cálculos en forma exacta (obviamente se puede usar calculadora).

194 $A_1 = \begin{bmatrix} 1 & -21 & 1/6 \\ -1 & -3 & \frac{215}{42} \\ 5 & -6 & -\frac{25}{42} \\ 3 & 2 & \frac{9}{14} \end{bmatrix}$

195 $A_2 = \begin{bmatrix} 2 & -2 & -2/7 \\ -4 & -10 & -2/21 \\ 5 & -19 & -\frac{29}{21} \\ 2 & 5 & 1/21 \end{bmatrix}$

196 Utilizar el resultado del ejercicio 194 para resolver el sistema $A_1\vec{x} = [\ 43 \quad 5 \quad 17 \quad -1\ ]^t$.

197 Utilizar el resultado del ejercicio 195 para resolver el sistema $A_2\vec{x} = [\ 6 \quad -292 \quad -265 \quad 146\ ]^t$.

En los ejercicios 198 a 201 reducir la matriz A a su forma similar superior de Hessenberg por el método dado en el apartado 8.4.3; es decir, encontrar una matriz ortogonal H tal que $Z = H^t AH$ esté en forma superior de Hessenberg. Realizar los cálculos sin el uso de programa alguno de computación y en forma exacta (obviamente puede auxiliarse de una calculadora)

198 $A = \begin{bmatrix} -1 & 2 & 1 \\ -3 & -1 & -1 \\ 4 & 3 & 1 \end{bmatrix}$

199 $A = \begin{bmatrix} 13 & 13 & -26 \\ -5 & 26 & 0 \\ 12 & 39 & 0 \end{bmatrix}$

200 $A = \begin{bmatrix} -3 & 1 & 3 \\ 8 & -2 & -1 \\ -6 & 1 & 2 \end{bmatrix}$

201 $A = \begin{bmatrix} 17 & -17 & 34 \\ -15 & -34 & 17 \\ 8 & -17 & -34 \end{bmatrix}.$

En los ejercicios 202 a 205 encontrar la reducción a forma superior de Hessenberg de la matriz indicada; trabajar con calculadora o utilizar la función `householder1` (cfr. figura 8-17) en MATLAB como se hizo en el ejemplo 8.47; presentar los resultados redondeados a 4 cifras decimales.

202 $A = \begin{bmatrix} 1 & -1 & -3 & -4 \\ 2 & 1 & 2 & 2 \\ 1 & 2 & 1 & 1 \\ 2 & 3 & 1 & 2 \end{bmatrix}.$

203 $A = \begin{bmatrix} -3 & 1 & -3 & -2 \\ 4 & -1 & 2 & 1 \\ -2 & 1 & 1 & 2 \\ 4 & 2 & 1 & 1 \end{bmatrix}.$

204 $A = \begin{bmatrix} 1 & -1 & 2 & -1 & 1 \\ 1 & 2 & -3 & 1 & -2 \\ -5 & 2 & -4 & 1 & -2 \\ 3 & -1 & 1 & 1 & 2 \\ 1 & -1 & 1 & -1 & 5 \end{bmatrix}.$

205 $A = \begin{bmatrix} -3 & 2 & 1 & -1 & 4 \\ 4 & -3 & 2 & 1 & -1 \\ -2 & 2 & -1 & 1 & 3 \\ 5 & 1 & -1 & 1 & 2 \\ 2 & 2 & -2 & 1 & 1 \end{bmatrix}.$

En los ejercicios 206 a 215 se recomienda utilizar, cuando así convenga, las matrices de rotación y de Givens definidas en (8.59) y (8.60), respectivamente; y las relaciones (8.63) y (8.64) del apartado 8.4.4.

206 Hallar la matriz de rotación y el vector que resulta de rotar $\vec{u} = (1, 1)$ treinta grados conservando su norma.

207 Hallar la matriz de rotación y el vector que resulta de rotar $\vec{u} = (1, 1)$ cuarenta y cinco grados conservando su norma.

208 Hallar la matriz de reflexión y el vector que resulta de reflejar $\vec{u} = (3, 1)$ sobre la recta $x = 0$ conservando su norma.

209 Hallar la matriz de reflexión y el vector que resulta de reflejar $\vec{u} = (3, 1)$ sobre la recta $y = x$ conservando su norma.

210 Encontrar la matriz de reflexión que al vector $\vec{x} = (-1, 4, 3)$ lo transforme en un vector cuya tercera componente sea cero, deje invariante la primera coordenada y preserve la norma.

211 Encontrar la matriz de reflexión que al vector $\vec{x} = (2, 3, -4, 1)$ lo transforme en un vector cuya tercera componente sea cero, deje invariante la primera y cuarta coordenadas y preserve la norma.

212 Encontrar la matriz de rotación que al vector $\vec{x} = (2, -3, -4, -1)$ lo transforme en un vector cuya tercera componente sea cero, deje invariante la primera y cuarta coordenadas y preserve la norma.

213 Si

$$A = \begin{bmatrix} 195 & 432 & -3445 \\ 260 & -74 & 9490 \\ 780 & 480 & 2275 \end{bmatrix}$$

encontrar una factorización QR, $A = QR$, donde R sea una matriz triangular superior y Q una matriz ortogonal que sea producto de matrices de Givens.

214 Demostrar el siguiente teorema:

Teorema: **[Factorización QR de Givens]** Sea $A \in \mathfrak{M}_{m\times n}(\mathbb{R})$, con $m \geq n$. Entonces existen dos matrices: $Q \in \mathfrak{M}_m(\mathbb{R})$ —matriz ortogonal, producto de matrices de Givens— y $\widehat{R} \in \mathfrak{M}_n(\mathbb{R})$ —matriz triangular superior— tales que

$$A = QR$$

donde $R = \begin{bmatrix} \widehat{R} \\ \mathscr{O}_{m-n,n} \end{bmatrix}$.

215 Utilizar el ejercicio precedente para hallar una factorización QR de Givens de la matriz

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 3 \\ 2 & 2 & -2 \\ 3 & 1 & 1 \end{bmatrix}.$$

Aproximación de valores y vectores propios (respuestas en páginas 1103-1104)

216 **(Teorema de Gershgorin)** Sea $A = [a_{ij}]$ una matriz cuadrada de orden n (con entradas reales y/o complejas) y sea

$$D = \bigcup_{i=1}^{n} D_i$$

donde

$$D_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \right\}$$

para cada $i = 1, 2, \ldots, n$. Demostrar que entonces

$$\lambda \in D$$

para todo valor propio λ de A. A los conjuntos D_i se les dice discos de Gershgorin.

En los ejercicios 217 a 220 encontrar los discos de Gershgorin, definidos en el ejercicio anterior, para la matriz A ahí dada y encontrar α, β tales que $\alpha \leq |\lambda| \leq \beta$ para todo valor propio de A.

217 $A = \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix}.$

218 $A = \begin{bmatrix} 1+i & 1 & -3 \\ 2 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix}.$

219 $A = \begin{bmatrix} 1+3i & \frac{1}{4}i & 1/2 \\ 1 & -2 & -1/4 \\ 1 & -1 & 5 \end{bmatrix}$.

220 $A = \begin{bmatrix} 2-3i & \frac{1}{2}i & 1/2 \\ 1-i & -2 & -1/4 \\ 3+21i & -2 & 1 \end{bmatrix}$.

221 Sea $A \in \mathfrak{M}_n(\mathbb{R})$ una matriz con valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$.

(a) Probar que $\operatorname{tra}(A^m) = \lambda_1^m + \lambda_2^m + \cdots + \lambda_n^m$ para todo $m = 1, 2, \ldots$.

(b) Suponer que $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$, mostrar que

$$\lambda_1 = \lim_{m\to\infty} \frac{\operatorname{tra}(A^{m+1})}{\operatorname{tra}(A^m)}$$

En los ejercicios 222 a 228 calcular las iteraciones indicadas, μ_k y $\vec{u}_k$, mediante la relación de recurrencia (8.66) del método de la potencia contenido en el teorema 8.11 (pág. 926), para calcular las aproximaciones del valor propio dominante y de un vector propio correspondiente de la matriz A tomando como valor inicial al vector $\vec{u}_0$ y el número entero k ahí dados.

222 $A = \begin{bmatrix} -1 & 2 & 12 \\ 1 & 0 & -6 \\ 2 & 4 & 4 \end{bmatrix}$, $\vec{u}_0 = (1, 1, 1)$, $k = 4$.

223 $A = \begin{bmatrix} 9 & 14 & 6 \\ -4 & -6 & -3 \\ 3 & 6 & 5 \end{bmatrix}$, $\vec{u}_0 = (1, 1, 2)$, $k = 4$.

224 $A = \begin{bmatrix} -3 & -4 & 12 \\ 3 & 5 & -6 \\ 1 & 2 & 5 \end{bmatrix}$, $\vec{u}_0 = (1, 2, 2)$, $k = 5$.

225 $A = \begin{bmatrix} -3 & -4 & -6 \\ 0 & -1 & 3 \\ -2 & -4 & -4 \end{bmatrix}$, $\vec{u}_0 = (1, 1, 3)$, $k = 3$.

226 $A = \begin{bmatrix} -3 & -4 & -6 \\ 0 & -1 & 3 \\ -2 & -4 & -4 \end{bmatrix}$, $\vec{u}_0 = (1, -0.4, -0.3)$, $k = 3$.

227 $A = \begin{bmatrix} -4 & -10 & -12 \\ 1 & 3 & 6 \\ -4 & -8 & -5 \end{bmatrix}$, $\vec{u}_0 = (1, 1, 2)$, $k = 4$.

228 $A = \begin{bmatrix} -10 & -18 & -10 \\ 4 & 7 & 5 \\ -5 & -10 & -6 \end{bmatrix}$, $\vec{u}_0 = (1, 1, 1)$, $k = 3$.

229 Sea $\|\cdot\|$ la norma euclidiana en $\mathbb{R}^n$. Demostrar que en el método de la potencia del teorema 8.11 la sucesión $(\vec{u}_k)$, definida por la relación de recurrencia (8.66), satisface

$$\|\vec{u}_k - c\vec{v}_1\| \leq M\left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k$$

para k suficientemente grande, donde $c = \text{máxc}(\vec{v}_1)$, $M = \sum_{j=2}^{n} |\beta_2| \|v_j\|$ y $\beta_j = \alpha_j/\text{máxc}(\alpha_1\vec{v}_1)$. Lo cual demuestra que la velocidad de convergencia de la sucesión $(\vec{u}_k)$ a un vector propio es de orden $(|\lambda_2|/|\lambda_1|)^k$.

230 **(Cociente de Rayleigh)** Considere que A es una matriz simétrica real. Se define la función $r_A : \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\} \to \mathbb{R}$ como

$$r_A(\vec{x}) = \frac{A\vec{x} \cdot x}{\vec{x} \cdot \vec{x}} \qquad \text{(R)}$$

donde $\vec{u} \cdot \vec{v}$ es el producto punto en $\mathbb{R}^n$.

(a) Mostrar que para todo $i = 1, 2, \ldots, n$,

$$\frac{\partial r_A}{\partial x_i}(\vec{x}) = \frac{2}{\vec{x} \cdot \vec{x}}(A\vec{x} - r(\vec{x})\vec{x})_i$$

(se utiliza, como en otros casos, la notación $(\vec{u})_i$ para la coordenada i del vector $\vec{u}$).

(b) Probar que $\vec{u}$ es un vector propio de la matriz A si y sólo si $\vec{u}$ es un punto crítico de la función r_A.

A la función r_A se le llama **cociente de Rayleigh** para la matriz A.

231 Sean A una matriz real simétrica, λ un valor propio de A y $\vec{u}$ un vector propio correspondiente a λ. Mostrar que si r_A es la función cociente de Rayleigh definida por la relación (R) del ejercicio anterior, entonces $r_A(\vec{u}) = \lambda$.

232 Sea $A \in \mathfrak{M}_n(\mathbb{R})$ una matriz simétrica; por tanto, los valores propios de A, $\lambda_1, \lambda_2, \ldots, \lambda_n$, son todos reales y $\mathbb{R}^n$ tiene una base ortonormal de vectores propios correspondientes, $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, de sendos valores propios. Demostrar que existe una constante K tal que

$$|r_A(\vec{x}) - r_A(\vec{v}_j)| \leq K \|\vec{x} - \vec{v}_j\|^2.$$

para todo $\vec{x} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$. De esta manera, si $\vec{x}$ es un vector cercano a uno de los vectores propios ortonormales, entonces el cociente de Rayleigh en ese vector es cercano a un valor propio de la matriz A. En el siguiente ejercicio se hace patente este hecho.

233 **(Método de la potencia con cociente de Rayleigh)** Sea $A \in \mathfrak{M}_n(\mathbb{R})$ una matriz simétrica con valores propios (reales) $\lambda_1, \lambda_2, \ldots, \lambda_n$, y vectores propios correspondientes $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, respectivamente, que forman una base ortonormal de $\mathbb{R}^n$; se supone que λ_1 es un valor propio dominante. Se definen, por recurrencia, las siguientes sucesiones:

$$\left.\begin{aligned} \vec{z}_k &= A\vec{u}_{k-1} \\ \vec{u}_k &= \frac{1}{\text{máxc}(\vec{z}_k)}\vec{z}_k \\ \mu_k &= r_A(\vec{u}_k) \end{aligned}\right\} \qquad \text{(R2)}$$

donde r_A es el cociente de Rayleigh definido por la relación (R) del ejercicio 230 y $\vec{u}_0$ es una aproximación inicial para un vector propio correspondiante a λ_1 ($\vec{u}_0$ puede ser el resultado de varias iteraciones del método de la potencia). Demostrar que:

(a) Existe c tal que $\lim_{k\to\infty} \vec{u}_k = c\vec{v}_1$

(b) $\lim_{k\to\infty} \mu_k = \lambda_1$

234 Mostrar que si se tienen las condiciones del ejercicio anterior, entonces existen dos constantes C_1 y C_2 tales que:

(a) $\|\vec{u}_k - c\vec{v}_1\| \leq C_1 \left|\dfrac{\lambda_2}{\lambda_1}\right|^k$

(b) $|\mu_k - \lambda_1| \leq C_2 \left|\dfrac{\lambda_2}{\lambda_1}\right|^{2k}$

para k suficientemente grande; es decir, las velocidades de convergencia de $(\vec{u}_k)$ y (μ_k) son, respectivamente, de órdenes lineal y cuadrática.

En los ejercicios 235 a 240 aplicar las relaciones de recurrencia (R2), del método de la potencia con cociente de Rayleigh expuesto en el ejercicio 233, a la matriz A para calcular las iteraciones $\vec{u}_k$ y μ_k para el número k indicado y el vector inicial $\vec{u}_0$ dado.

235 $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & 3 \end{bmatrix}$, $\vec{u}_0 = (1, 1, 1)$, $k = 2$.

236 $A = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}$, $\vec{u}_0 = (1, 1, 2)$, $k = 3$.

237 $A = \begin{bmatrix} 4 & 4 & 0 \\ 4 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}$, $\vec{u}_0 = (1, 1/2, 0)$, $k = 2$.

238 $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & -2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, $\vec{u}_0 = (-1/2, -1, 1/2)$, $k = 2$.

239 $A = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 1 & 3 \\ 3 & 3 & 2 \end{bmatrix}$, $\vec{u}_0 = (1/2, 1/2, 1)$, $k = 2$.

240 $A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & -3 & 2 \\ -1 & 2 & 4 \end{bmatrix}$, $\vec{u}_0 = (-1/2, 1/2, 1)$, $k = 2$.

241 Construir un programa en MATLAB, o en algún otro lenguaje de programación, para implementar el método de la potencia con cociente de Rayleigh del ejercicio 233. Las entradas deben ser una matriz cuadrada simétrica, un vector inicial, un margen de tolerancia entre las diferencias de iteraciones sucesivas y el número máximo de iteraciones permisible. La salida debe contener las aproximaciones del valor

propio dominante y del vector propio correspondiente. Además, el programa debe desplegar mensajes si la matriz de entrada no es simétrica y si se rebasa el número máximo de iteraciones permitido.

En los ejercicios 242 a 251 utilizar las relaciones (8.70), página 932, del método de deflación del apartado 8.5.2 para hallar los valores propios de la matriz A si ésta tiene el valor propio dominante λ_1 y el vector propio correspondiente $\vec{v}_1$ que se indican.

242 $A = \begin{bmatrix} -4 & -6 & -4 \\ 1 & 1 & 2 \\ -2 & -4 & -3 \end{bmatrix}$, $\lambda_1 = -3$, $\vec{v}_1 = (2, -1, 1)$.

243 $A = \begin{bmatrix} -4 & -6 & -6 \\ 1 & 1 & 3 \\ -2 & -4 & -4 \end{bmatrix}$, $\lambda_1 = -4$, $\vec{v}_1 = (-2, 1, -1)$.

244 $A = \begin{bmatrix} 9 & -2 & -12 \\ 7 & 0 & -6 \\ 2 & -2 & -5 \end{bmatrix}$, $\lambda_1 = 7$, $\vec{v}_1 = (1, 1, 0)$.

245 $A = \begin{bmatrix} 7 & -2 & -11 \\ 6 & -1 & -5 \\ 2 & -2 & -6 \end{bmatrix}$, $\lambda_1 = 5$, $\vec{v} = (1, 1, 0)$.

246 $A = \begin{bmatrix} 2 & 6 & -4 \\ -4 & -9 & 2 \\ -1 & -2 & -3 \end{bmatrix}$, $\lambda_1 = -6$, $\vec{v}_1 = (1, -2, -1)$.

247 $A = \begin{bmatrix} 1 & 10 & 22 \\ -2 & -8 & -11 \\ 4 & 8 & 7 \end{bmatrix}$, $\lambda_1 = 7$, $\vec{v}_1 = (2, -1, 1)$.

248 $A = \begin{bmatrix} -1 & 15 & -2 & 10 \\ -2 & -8 & 0 & -6 \\ 5 & -5 & 2 & 0 \\ -1 & -5 & -1 & 1 \end{bmatrix}$, $\lambda_1 = -4$, $\vec{v}_1 = (-5, 1, 5, 1)$.

249 $A = \begin{bmatrix} -1 & 19 & -7 & 31 \\ 0 & -6 & 2 & -8 \\ 4 & -7 & 6 & -23 \\ 2 & 1 & 2 & -5 \end{bmatrix}$, $\lambda_1 = -4$, $\vec{v}_1 = (5, -1, -5, -1)$.

250 $A = \begin{bmatrix} -1 & 16 & 1 & 12 \\ -10 & -17 & 1 & -6 \\ 4 & 32 & 2 & 18 \\ -3 & 16 & 1 & 14 \end{bmatrix}$, $\lambda_1 = -3$, $\vec{v}_1 = (1, -1, 2, 1)$.

251 $A = \begin{bmatrix} 2 & 8 & -2 & 5 \\ -19 & -15 & 4 & 1 \\ 5 & 16 & -3 & 7 \\ 0 & 8 & -2 & 7 \end{bmatrix}$, $\lambda_1 = -7$, $\vec{v}_1 = (1, -5/4, 2, 1)$.

En los ejercicios 252 a 257 hacer lo siguiente:

(a) Aplicar el método de la potencia del teorema 8.11 a la matriz A, con el vector inicial $\vec{u}_0$, para obtener las primeras k iteraciones μ_k y $\vec{u}_k$ y aproximar el valor propio dominante, λ_1, y el vector propio correspondiente, $\vec{v}_1$, de la matriz A.

(b) Redondear los valores obtenidos en el primer inciso a una cifra decimal significativa y comprobar que $\lambda_1 = \mu_k$ es, efectivamente, un valor propio de A con vector propio correspondiente $\vec{v}_1 = \vec{u}_k$.

(c) Utilizar el inciso *(b)* para, con el método de deflación, encontrar los restantes valores propios de la matriz A.

252 $k = 2$, $A = \begin{bmatrix} -1 & -6 & -4 \\ -11 & 4 & 14 \\ 6 & -6 & -11 \end{bmatrix}$, $\vec{u}_0 = (-1/2, -1/2, 0)$.

253 $k = 3$, $A = \begin{bmatrix} 5 & -4 & -10 \\ 8 & 1 & -6 \\ 0 & -4 & -5 \end{bmatrix}$, $\vec{u}_0 = (0.2, -1, 0.38)$.

254 $k = 3$, $A = \begin{bmatrix} -3 & -6 & -2 \\ -13 & 4 & 16 \\ 6 & -6 & -11 \end{bmatrix}$, $\vec{u}_0 = (1/4, 1/3, 2/3)$.

255 $k = 2$, $A = \begin{bmatrix} -2 & -10 & -6 \\ -19 & 7 & 24 \\ 10 & -10 & -18 \end{bmatrix}$, $\vec{u}_0 = (1/2, 1/2, 0)$.

256 $k = 3$, $A = \begin{bmatrix} -6 & -8 & -1 \\ -19 & 5 & 23 \\ 8 & -8 & -15 \end{bmatrix}$, $\vec{u}_0 = (-3, -3, 1)$.

257 $k = 2$, $A = \begin{bmatrix} -7 & -15 & -6 \\ -23 & 12 & 33 \\ 8 & -15 & -21 \end{bmatrix}$, $\vec{u}_0 = (1, -4/5, 1)$.

En los ejercicios 258 a 261, encontrar los valores propios de la matriz A por medio del método de deflación del apartado 8.5.2 siguiendo el procedimiento que se llevó a cabo en el ejemplo 8.60.

258 $A = \begin{bmatrix} 50 & -53 & 46 & -49 \\ 12 & -15 & 12 & -12 \\ -7 & 7 & -5 & 7 \\ 22 & -22 & 20 & -21 \end{bmatrix}$.

259 $A = \begin{bmatrix} -17 & 15 & -12 & 27 \\ 3 & -5 & 3 & -3 \\ 4 & -4 & 9 & -4 \\ -17 & 17 & -7 & 27 \end{bmatrix}$.

260 $A = \begin{bmatrix} -12 & 4 & 31 & 35 \\ -3 & 0 & -4 & 5 \\ -1 & 2 & -9 & -2 \\ -8 & 5 & 26 & 24 \end{bmatrix}$.

261 $A = \begin{bmatrix} -15 & 16 & 16 & 32 \\ -12 & 18 & -13 & 14 \\ -1 & 2 & -9 & -2 \\ -2 & -1 & 20 & 12 \end{bmatrix}$.

262 Sean $A \in \mathfrak{M}_n(\mathbb{R})$ y λ un valor propio de A con un vector propio correspondiente $\vec{v}$. Mostrar que si ρ es un escalar, entonces $\lambda - \rho$ es un valor propio de la matriz $A - \rho I_n$ con $\vec{v}$ un vector propio correspondiente.

263 Sean $A \in \mathfrak{M}_n(\mathbb{R})$, ρ un escalar y μ un valor propio de la matriz $A - \rho I_n$ con $\vec{v}$ un vector propio correspondiente. Mostrar que $\lambda = \mu + \rho$ es un valor propio de la matriz A con $\vec{v}$ un vector propio correspondiente.

264 Sea $A \in \mathfrak{M}_n(\mathbb{R})$ una matriz que es no singular (invertible) y λ un valor propio de A distinto de cero con $\vec{v}$ un vector propio correspondiente. Demostrar que $\mu = \lambda^{-1}$ es un valor propio de A^{-1} con $\vec{v}$ un vector propio correspondiente.

En los ejercicios 265 a 274 λ es un valor propio de la matriz A de multiplicidad algebraica uno. *(a)* Calcular, con el método de iteración inversa del teorema 8.12, un vector propio $\vec{u}$ correspondiente al valor propio λ utilizando como vector inicial al vector $\vec{u}_0$ dado. *(b)* Comprobar que, efectivamente, el vector $\vec{u}$ es un vector propio correspondiente al valor propio λ efectuando los productos $A\vec{u}$ y $\lambda\vec{u}$.

265 $A = \begin{bmatrix} -3 & -7 & 11 \\ -1 & -3 & 3 \\ 0 & 3 & -6 \end{bmatrix}, \lambda = -2, \vec{u}_0 = (1, 1, 1).$

266 $A = \begin{bmatrix} -3 & -7 & 11 \\ -1 & -3 & 3 \\ 0 & 3 & -6 \end{bmatrix}, \lambda = -9, \vec{u}_0 = (1, 1, 1)$

267 $A = \begin{bmatrix} 2 & -12 & 6 \\ 4 & -8 & 8 \\ 5 & -2 & 9 \end{bmatrix}, \lambda = 8, \vec{u}_0 = (-1, 2, 2).$

268 $A = \begin{bmatrix} 2 & -12 & 15 \\ 1 & -5 & 5 \\ -1 & 4 & -6 \end{bmatrix}, \lambda = -7, \vec{u}_0 = (1, 1/2, 1/2).$

269 $A = \begin{bmatrix} -4 & -6 & 9 \\ -1 & -3 & 3 \\ 1 & 2 & -4 \end{bmatrix}, \lambda = -9, \vec{u}_0 = (1, 1/2, -1/2).$

270 $A = \begin{bmatrix} -8 & 4 & -8 \\ -14 & 4 & -10 \\ 0 & -3 & 5 \end{bmatrix}, \lambda = -1, \vec{u}_0 = (1, 1, 1).$

271 $A = \begin{bmatrix} -8 & 4 & -8 \\ -14 & 4 & -10 \\ 0 & -3 & 5 \end{bmatrix}, \lambda = 6, \vec{u}_0 = (1, 1, 1).$

272 $A = \begin{bmatrix} 2 & -12 & 6 \\ 4 & -8 & 8 \\ 5 & -2 & 9 \end{bmatrix}, \lambda = -1, \vec{u}_0 = (1, 1, 1).$

273 $A = \begin{bmatrix} -3 & -1 & 2 \\ 16 & 8 & 2 \\ -9 & -6 & -7 \end{bmatrix}, \lambda = -5, \vec{u}_0 = (1, 1, 1).$

274 $A = \begin{bmatrix} -3 & -1 & 2 \\ 16 & 8 & 2 \\ -9 & -6 & -7 \end{bmatrix}, \lambda = 1, \vec{u}_0 = (1, 1, 1).$

En el ejercicio 233, pág. 974, se mostró que el método de la potencia se puede combinar con el cociente de Rayleigh (cfr. ejercicio 230) para aumentar la velocidad de convergencia al valor propio dominante; sin embargo, la velocidad de convergencia al vector propio correspondiente aún es lineal en este método (cfr. ejercicio 234, pág. 975); para acelerar la convergencia es posible utilizar conjuntamente el método de iteración inversa con el cociente de Rayleigh. El algoritmo se plantea de la siguiente manera:

Iteración cociente de Rayleigh: Sea A una matriz simétrica con valores propios λ_i y vectores propios correspondientes $\vec{v}_i$, $i = 1, 2, \ldots, n$, respectivamente, que forman una base ortonormal del espacio $\mathbb{R}^n$. Si el vector propio λ_1 es dominante, $|\lambda_1| > |\lambda_2| \geq \cdots |\lambda_n|$, $\vec{u}_0$ es un vector inicial dado y se definen, en forma recurrente, las siguientes sucesiones

$$\left.\begin{aligned} \rho_k &= \frac{A\vec{u}_k \cdot \vec{u}_k}{\vec{u}_k \cdot \vec{u}_k} \\ (A - \rho_k I_n)\vec{z}_{k+1} &= \vec{u}_k \\ \vec{u}_{k+1} &= \frac{1}{\text{máxc}(\vec{z}_{k+1})}\vec{z}_{k+1} \end{aligned}\right\} \qquad \text{(RI)}$$

para $k = 0, 1, 2, \ldots$, entonces

$$\lim_{k\to\infty} \rho_k = \lambda_1 \quad \text{y} \quad \lim_{k\to\infty} \vec{u}_k = c\vec{v}_1$$

para alguna constante c. Conviene obtener la solución del sistema $(A - \rho_k I_n)\vec{z}_{k+1} = \vec{u}_k$, en cada iteración, por el método de Gauss con la estrategia de pivote parcial.

En los ejercicios 275 a 284 aplicar las relaciones de recurrencia (RI) del método iterativo del cociente de Rayleigh para aproximar el valor propio dominante (λ_1) y un vector propio correspondiente ($\vec{x}_1$) de la matriz simétrica A. El valor inicial $\vec{u}_0$ del ciclo (RI) se determina aplicando la primera iteración del método de la potencia con vector inicial, aquel que tiene todas sus coordenadas iguales a la unidad; después de cada iteración se debe calcular la diferencia $A\vec{u}_k - \rho_k\vec{u}_k$ para decidir si ésta es suficientemente cercana a cero para detener el proceso.

275 $A = \begin{bmatrix} 3 & 1 & 4 \\ 1 & 3 & 2 \\ 4 & 2 & -2 \end{bmatrix}$.

276 $A = \begin{bmatrix} 3 & 1 & 4 \\ 1 & 3 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

277 $A = \begin{bmatrix} 1 & 1 & 4 \\ 1 & 2 & 1 \\ 4 & 1 & 3 \end{bmatrix}$.

278 $A = \begin{bmatrix} 7 & 5 & 1 \\ 5 & 1 & 3 \\ 1 & 3 & 2 \end{bmatrix}$.

279 $A = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 3 & 7 \end{bmatrix}$.

280 $A = \begin{bmatrix} 8 & 3 & 2 \\ 3 & 7 & 4 \\ 2 & 4 & -3 \end{bmatrix}$.

281 $A = \begin{bmatrix} 4 & 1 & 2 & 3 \\ 1 & 4 & 1 & 2 \\ 2 & 1 & 4 & 1 \\ 3 & 2 & 1 & 4 \end{bmatrix}$.

282 $A = \begin{bmatrix} 6 & 1 & 2 & 3 \\ 1 & 6 & 1 & 2 \\ 2 & 1 & 4 & 1 \\ 3 & 2 & 1 & 4 \end{bmatrix}$

283 $A = \begin{bmatrix} 3 & 1 & -1 & 2 \\ 1 & -3 & 1 & 3 \\ -1 & 1 & 6 & 1 \\ 2 & 3 & 1 & 8 \end{bmatrix}.$

284 $A = \begin{bmatrix} 1 & 1 & -1 & 3 \\ 1 & -1 & 2 & 3 \\ -1 & 2 & 9 & 1 \\ 3 & 3 & 1 & 7 \end{bmatrix}.$

285 Construir en MATLAB (o en algún otro lenguaje) un programa para implementar el método de iteración del cociente de Rayleigh, relaciones (RI), que tenga como criterio para detenerse un margen de tolerancia entre la norma cúbica de los vectores $A\vec{u}_k - \rho_k\vec{u}_k$, si A es la matriz de entrada; el programa debe emitir un mensaje adecuado para no rebasar un número máximo de iteraciones especificado por el usuario el cual, también, determina el margen de tolerancia. La aproximación inicial, $\vec{u}_0$, debe ser la primera estimación que resulta del método de la potencia. El programa debe incluir los correspondientes mensajes si el usuario introduce alguna matriz que no sea simétrica.

En los ejercicios 286 a 291 aplicar el método QR a la matriz A hasta obtener una matriz de la forma 8.73 y, entonces, calcular los valores propios de los bloques B_j para estimar los valores propios de la matriz A como se hizo en el ejemplo 8.64 de la página 940. No utilizar programa alguno; hacer las operaciones en forma aritmética con calculadora.

286 $A = \begin{bmatrix} 7.0000 & 2.0083 & 2.5529 \\ 0.0001 & 2.0425 & -2.2891 \\ 0.0000 & 1.7668 & 0.9575 \end{bmatrix}.$

287 $A = \begin{bmatrix} 9.0011 & -11.4325 & 52.6826 \\ 0.0001 & 0.1666 & 1.8596 \\ -0.0002 & -0.5525 & -0.1676 \end{bmatrix}.$

288 $A = \begin{bmatrix} 11.0000 & -0.6395 & 11.5021 & -61.1011 & 49.7563 \\ 0.0001 & -0.5000 & -2.1408 & 8.1645 & -4.7431 \\ 0.0000 & 0.5839 & 0.5000 & -4.0043 & 6.0313 \\ 0.0000 & 0.0000 & 0.0000 & -1.0000 & 1.2910 \\ 0.0000 & 0.0000 & 0.0000 & -0.7746 & 0.0000 \end{bmatrix}.$

289 $A = \begin{bmatrix} -6.9994 & 33.9395 & -8.5347 \\ 0.0001 & -1.0000 & -1.5071 \\ -0.0002 & 3.3181 & 0.9994 \end{bmatrix}.$

290 $A = \begin{bmatrix} 16.0019 & -18.7857 & -31.5939 & -238.2664 \\ -0.0004 & -1.9996 & -5.6865 & 0.6804 \\ 0.0000 & 0.0000 & -2.8113 & -3.5072 \\ 0.0002 & -0.0002 & 2.7537 & -1.1910 \end{bmatrix}.$

291 $A = \begin{bmatrix} 8.9985 & 5.1170 & -67.8826 \\ 0.0002 & 3.0001 & 3.3151 \\ -0.0003 & 0.9044 & -2.9986 \end{bmatrix}.$

En los ejercicios 292 a 295 aplicar el método QR para producir, en cada parte del proceso, matrices similares a la matriz A de la forma

$$\begin{bmatrix} \lambda_k & b_{k,k+1} & \cdots & b_n \\ 0 & & & \\ \vdots & & \Gamma_k & \\ 0 & & & \end{bmatrix}$$

hasta obtener los valores propios de ésta como se hizo en el ejemplo 8.66 de la página 942. No utilizar programa alguno y hacer las operaciones en forma aritmética con calculadora.

292 $A = \begin{bmatrix} 7.4289 & -1.2973 & -0.5584 \\ -0.0002 & -3.1424 & -0.0753 \\ 0.0000 & 0.0003 & 1.7134 \end{bmatrix}.$

293 $A = \begin{bmatrix} 11.9213 & 2.2919 & -0.1849 \\ -0.0001 & 3.4608 & -0.8422 \\ 0.0000 & 0.0021 & -1.3821 \end{bmatrix}.$

294 $A = \begin{bmatrix} 11.9217 & 13.2998 & -12.8920 \\ -0.0003 & 3.4582 & 20.4432 \\ 0.0000 & 0.0004 & -1.3799 \end{bmatrix}.$

295 $A = \begin{bmatrix} -6.5250 & -5.0777 & 3.8653 \\ 0.0001 & 3.5250 & -0.5248 \\ 0.0000 & -0.0002 & -2.0000 \end{bmatrix}.$

En los ejercicios 296 a 301 utilizar las funciones `reduccion_hhh` y `ciclo_QR` de las figuras 8-19 (pág. 916) y 8-26 (pág. 940), respectivamente, o programas similares en algún lenguaje que el lector haya construido, para calcular los valores propios de la matriz A por medio del método QR con redución de Hessenberg, algoritmo 8.11, como se hizo en el ejemplo 8.70 de la página 946. Realizar el menor número posible de iteraciones.

296 $A = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 1 & 1 \\ 5 & 3 & 4 \end{bmatrix}.$

297 $A = \begin{bmatrix} 1 & 4 & 1 \\ 5 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}.$

298 $A = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 1 & 1 \\ 5 & -3 & 3 \end{bmatrix}.$

299 $A = \begin{bmatrix} 21 & -24 & 24 & -16 \\ 8 & -11 & 8 & -8 \\ -1 & 1 & -3 & -1 \\ 7 & -7 & 8 & -4 \end{bmatrix}.$

300 $A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 1 & 2 & 1 & 1 \\ 2 & 1 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}.$

301 $A = \begin{bmatrix} 4 & 3 & 7 & 1 \\ 4 & -5 & 9 & -1 \\ 0 & -1 & -1 & 1 \\ -2 & -2 & 1 & -5 \end{bmatrix}.$

En los ejercicios 302 a 307, aplicar el método QR con reducción de Hessenberg y desplazamientos, algoritmo 8.12, para encontrar los valores propios de la matriz A como se hizo en el ejemplo 8.71 de la página 947.

302 $A = \begin{bmatrix} 1 & 4 & 3 \\ 7 & 3 & 2 \\ 1 & 2 & 6 \end{bmatrix}.$

303 A es la matriz del ejercicio 296.

304 $A = \begin{bmatrix} 2 & 1 & 3 \\ 2 & 1 & 5 \\ 1 & 3 & 4 \end{bmatrix}.$

305 A es la matriz del ejercicio 298.

306 $A = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & 3 & 2 \\ 1 & 2 & -9 & 4 \\ 1 & 3 & 2 & 1 \end{bmatrix}.$

307 A es la matriz del ejercicio 301.

308 Demostrar que el polinomio característico de la matriz de Frobenius

$$A = \begin{bmatrix} -a_{n-1} & -a_{n-2} & -a_{n-3} & \cdots & -a_1 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \qquad \text{(F)}$$

está dado por

$$p_A(\lambda) = (-1)^n(\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + a_{n-3}\lambda^{n-3} \cdots + a_1\lambda + a_0)$$

Inversamente, probar que si

$$p(x) = \alpha_n x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0$$

es un polinomio de grado n, entonces las raíces de este polinomio son los valores propios de la matriz de Frobenius (F), donde $a_k = \alpha_k/\alpha_n$, $k = 0, 1, \ldots, n-1$. A la matriz, así formada, se le llama matriz de Frobenius asociada al polinomio $p(x)$.

En los ejercicios 309 a 314 utilizar el ejercicio 308 para calcular los valores propios de la matriz de Frobenius asociada al polinomio $p(x)$ y encontrar así las raíces de éste.

309 $p(x) = x^3 - 5x^2 + 2x + 8$.

310 $p(x) = x^3 - x^2 - 17x - 15$.

311 $p(x) = 5x^3 - 4x^2 + 7x - 2$.

312 $p(x) = x^4 - 5x^3 - 2x + 1$.

313 $p(x) = 24x^4 - 96x^3 + 72x^2 - 16x + 1$.

314 $p(x) = x^4 + 6x^3 + 5x^2 - 4x - 2$.

En los ejercicios 315 a 324 hallar los valores y vectores propios de la matriz A por el método de Jacobi (algoritmo 8.13, pág. 953).

315 $A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 4 \end{bmatrix}$.

316 $A = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix}$.

317 $A = \begin{bmatrix} 3 & -2 & -6 \\ -2 & 3 & 1 \\ -6 & 1 & 9 \end{bmatrix}$.

318 $A = \begin{bmatrix} 2 & -3 & -1 \\ -3 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$.

319 $A = \begin{bmatrix} 2 & -1 & 1 & -1 \\ -1 & 2 & 1 & -1 \\ 1 & 1 & 2 & 1 \\ -1 & -1 & 1 & 2 \end{bmatrix}$.

320 $A = \begin{bmatrix} 4 & -1 & 1 & -1 \\ -1 & 4 & 1 & -1 \\ 1 & 1 & 4 & 1 \\ -1 & -1 & 1 & 4 \end{bmatrix}$.

321 $A = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 4 & 1 & -1 \\ 1 & 1 & 4 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$.

322 $A = \begin{bmatrix} 2 & -2 & 2 & -2 \\ -2 & 2 & 2 & -2 \\ 2 & 2 & 2 & 2 \\ -2 & -2 & 2 & 2 \end{bmatrix}$.

323 $A = \begin{bmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 2 \end{bmatrix}$.

324 $A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$.

A Conjuntos, demostraciones e inducción matemática

A.1 Conjuntos

En este apartado se introducen los conceptos elementales y notaciones básicas de la teoría de conjuntos.

A.1.1 Conjuntos, elementos y subconjuntos

Intuitivamente, una colección de objetos formada de acuerdo con cierta regla que permite siempre decidir si un objeto dado, cualquiera, pertenece o no a dicha colección, es un conjunto. Si A es un conjunto y a es un objeto que forma parte de él, se dice que a es un **elemento** de A o que a **pertenece** al conjunto A y se escribe

$$a \in A$$

y si el objeto a no es elemento del conjunto A, se escribe

$$a \notin A.$$

Usualmente los conjuntos se denotan por letras mayúsculas y sus elementos por letras minúsculas; aunque obviamente hay excepciones.

Los conjuntos se pueden formar de dos maneras:

1. Por **extensión**: Colocando entre llaves los elementos del conjunto y separando éstos por comas. Por ejemplo, el conjunto

$$B = \{1, \clubsuit, \heartsuit, c\}.$$

2. Por **condición**: Los elementos que forman el conjunto se describen por medio de una o varias condiciones que deben cumplir. El formato en este caso es:

$$\{x \mid p(x)\} \quad \text{o} \quad \{x : p(x)\}$$

donde $p(x)$ es la condición o condiciones que deben cumplir los objetos x para conformar el conjunto; mientras que la línea vertical y los dos puntos, después de la literal x, son abreviaciones de la frase "tales que". Por ejemplo, si se desea formar el conjunto de los números primos escribimos

$$E = \{x : x \text{ es número primo}\} \quad \text{o} \quad E = \{x \mid x \text{ es número primo}\}.$$

Por supuesto que el conjunto E, descrito arriba, también se puede escribir por extensión:

$$E = \{2, 3, 5, 7, 11, 13, \ldots\}$$

donde los tres puntos suspensivos representan la frase "así sucesivamente".

Entonces, si B y E son los dos conjuntos que se definieron arriba, $\clubsuit \in B$, $2 \notin B$, $29 \in E$ y $121 \notin E$, por ejemplo.

Subconjuntos

Definición A.1 *Sean A y B dos conjuntos, se dice que B es un **subconjunto** de A si todo elemento de B pertenece también al conjunto A. En tal caso se emplea la notación*

$$B \subset A \quad o \quad B \subseteq A$$

para simbolizar este concepto. También se acostumbra decir que B está contenido *en A o que A* contiene *al conjunto B.*

En la figura A-1 *(a)* se da un esquema gráfico, llamado diagrama de Venn, para representar esta idea ($B \subset A$). Naturalmente escribiremos $B \not\subset A$ (o $B \nsubseteq A$) para decir que B no es un subconjunto de A (B no está contenido en A); lo cual significa que al menos un elemento de B no pertenece al conjunto A. La figura A-1 *(b)* contiene el diagrama de Venn para este hecho.

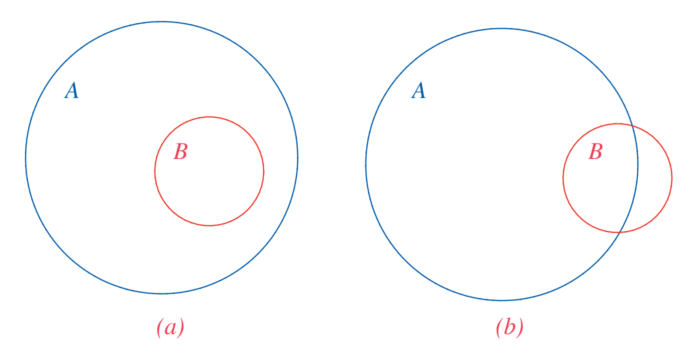


Figura A-1 • *(a)* Diagrama de Venn para representar $B \subset A$. *(b)* Diagrama de Venn para representar $B \not\subset A$.

En los diagramas de Venn se utilizan curvas planas simples y cerradas, por ejemplo circunferencias, cuyos puntos interiores representan a todos los elementos de un conjunto dado. Cuando se requiere se sombrean regiones entre curvas para, de esta manera, gráficamente dar un esquema de un conjunto.

► **Ejemplo A.1** Sean $A = \{2, 3, 5, 7, 11, 13, 17, 19, 23\}$, $B = \{3, 11, 17\}$ y $C = \{2, 3, 6, 7, 10\}$; entonces $B \subset A$, pero $C \not\subset A$.◄

Subconjuntos propios

De la definición A.1 se desprende que todo conjunto es subconjunto de sí mismo; es decir, $A \subset A$ para cualquier conjunto A. Si un conjunto B es subconjunto de un conjunto A, pero es distino de A, se dice

que B es un **subconjunto propio** de A y se escribe

$$B \subsetneq A$$

que significa $B \subset A$ y $B \neq A$. De esta manera, si A y B son los conjuntos del ejemplo A.1, entonces B es un subconjunto propio de A.

Igualdad de conjuntos

Definición A.2 *Dos conjuntos A y B son iguales, $A = B$, si y sólo si $A \subset B$ y $B \subset A$; es decir, todo elemento de A pertence a B y todo elemento de B pertenece a A.*

► **Ejemplo A.2** Sean $A = \{x \in \mathbb{R} \mid (x-2)(x-3) = 0\}$, $B = \{x \in \mathbb{R} \mid x^2 - 5x + 6 = 0\}$ y $C = \{2,3\}$; es claro que $A = B = C$.◄

► **Ejemplo A.3** Sean $A = \{5,4,1\}$ y $B = \{1,5,4\}$. Puesto que $A \subset B$ y $B \subset A$, se tiene $A = B$.◄

► **Ejemplo A.4** Si $A = \{1,2,3,1,4\}$ y $B = \{1,2,3,4\}$, ya que $A \subset B$ y $B \subset A$, entonces $A = B$.◄

❍ **Nota A.1** Observemos que, por la definición de igualdad de conjuntos, si se describe un conjunto por extensión, la repetición de un mismo elemento es redundante; por tanto, cuando listemos los elementos de un conjunto, separándolos por comas, es suficiente que cada elemento aparezca solamente una vez en la lista.

Conjunto universo

Cuando en un grupo de conjuntos A, B, C, etc., todos los elementos de cada uno de ellos pertenecen a un mismo conjunto $\mathscr{U}$; es decir, $A \subset \mathscr{U}$, $B \subset \mathscr{U}$, $C \subset \mathscr{U}$, etc., se dice que el conjunto $\mathscr{U}$ es un **conjunto universo** para esta colección de conjuntos. Por ejemplo, si en un contexto dado, todos los conjuntos son determinadas clases de números reales, un conjunto universo para ellos puede ser el conjunto $\mathbb{R}$ de todos los números reales. Con frecuencia no se especifica un conjunto universo y está dado implícitamente, pues en el contexto es claro cuál puede ser un conjunto universo. Así, para los conjuntos

$$\begin{aligned} A &= \{x \mid x \text{ fuma entre 1 y 3 cajetillas al día}\} \quad \text{y} \\ B &= \{x \mid x \text{ tiene memoria entre 40 y 80 gigabytes en disco duro}\} \end{aligned}$$

es obvio que pueden tener como conjuntos universo a $\mathscr{U}_1$, el conjunto de seres humanos, y a $\mathscr{U}_2$, el conjunto de computadoras, respectivamente.

Conjunto vacío

Es conveniente establecer un tipo especial de conjunto, el cual, por definición, no tiene elementos. A este conjunto se le llama **conjunto vacío** y se representa por el símbolo $\emptyset$. De la definición A.1, se desprende que $\emptyset \subset A$ para todo conjunto A; es decir, **el conjunto vacío es subconjunto de todo conjunto**; pues en caso contrario existiría un elemento $x \in \emptyset$ que no pertenece a cierto conjunto A, lo cual es una contradicción de la definición del conjunto vacío.

A.1.2 Operaciones con conjuntos

Unión de conjuntos

Definición A.3 *Sean A y B un par de conjuntos. Se define y denota la* ***unión*** *de ellos como*

$$A \cup B = \{x \mid x \in A \ o \ x \in B\}.$$

Esto es, la unión de dos conjuntos está conformada por aquellos elementos que pertenecen por lo menos a uno de ellos. La figura A-2 contiene el diagrama de Venn para esta operación de conjuntos.

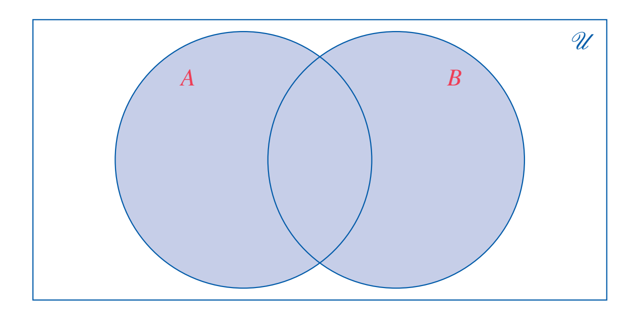


Figura A-2 • Diagrama de Venn para la unión de conjuntos. $A \cup B$: ▭

► **Ejemplo A.5** Si $A = \{4, s, w, 8, -4, d\}$ y $B = \{-1, t, u, w, \diamondsuit, -4, papa\}$, entonces

$$A \cup B = \{4, s, w, 8, -4, d, -1, t, u, \diamondsuit, papa\}.$$ ◄

Intersección de conjuntos

Definición A.4 *Si A y B son dos conjuntos, se define y denota la* ***intersección*** *de ellos como*

$$A \cap B = \{x : x \in A \ \ y \ x \in B\}.$$

Es decir, la intersección de dos conjuntos es un conjunto que está conformado por los elementos comunes (si los hay) de ambas colecciones. El diagrama de Venn para la intersección de dos conjuntos se encuentra contenido en la figura A-3

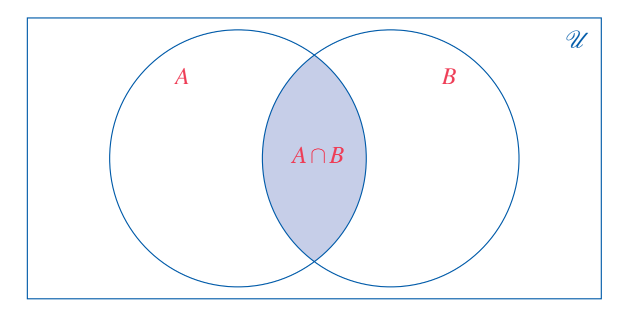


Figura A-3 • Diagrama de Venn para la intersección de dos conjuntos. $A \cap B$: ▭

▶ **Ejemplo A.6** Si A y B son los conjuntos del ejemplo A.5, entonces $A \cap B = \{-4, w\}$.◀

▶ **Ejemplo A.7** Sean $A = \{x \in \mathbb{N} \mid x \text{ es par}\}$, $B = \{x \in \mathbb{N} \mid x \text{ es impar}\}$, entonces $A \cap B = \emptyset$.◀

Definición A.5 *Si dos conjuntos tienen intersección vacía, se dice que son* ***disjuntos****.*

Complemento de un conjunto

Definición A.6 *Si A es un conjunto y $\mathscr{U}$ es un conjunto universo para A, se define y denota el complemento de A (relativo a $\mathscr{U}$) por*

$$A^c = \{x \in \mathscr{U} \mid x \notin A\}.$$

Esto es, A^c es el conjunto de elementos en el conjunto universo que no pertenecen al conjunto A. El diagrama de Venn para el complemento de un conjunto está contenido en la figura A-4.

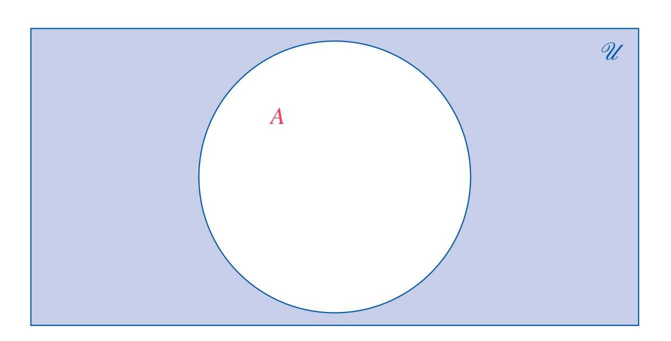


Figura A-4 • Diagrama de Venn para el complemento de un conjunto. A^c: ▭

▶ **Ejemplo A.8** Sean $\mathscr{U} = \mathbb{N} = \{1, 2, 3, \ldots\}$ y E el conjunto de números pares positivos, entonces $E^c = \{1, 3, 5, \ldots\}$.◀

Observe que $(A^c)^c = A$ para todo conjunto A.

Diferencia de conjuntos

Definición A.7 *Sean A, B dos conjuntos. La diferencia del conjunto A con el conjunto B se define y denota como*

$$A - B = \{x \mid x \in A \text{ y } x \notin B\}.$$

El diagrama de Venn para la diferencia de dos conjuntos se encuentra en la figura A-5.

Figura A-5 • Diagrama de Venn para la diferencia de conjuntos. $A - B$:

▶ **Ejemplo A.9** Si A y B son los conjuntos del ejemplo A.5, entonces

$$A - B = \{4, s, 8, d\} \text{ y}$$
$$B - A = \{-1, t, u, \diamondsuit, papa\}. ◀$$

Producto cartesiano de conjuntos

1. Sean a, b un par de objetos. Se denota la **pareja ordenada** formada por ellos como (a,b).
2. Dos parejas ordenadas (a,b) y (c,d) son iguales, $(a,b) = (c,d)$, si y sólo si $a = c$ y $b = d$.

De esta manera $(2,3) \neq (3,2)$; y de ahí el término usado para indicar el orden de los objetos en la frase *pareja ordenada* del primer inciso.

❍ **Nota A.2** Rigurosamente se define $(a,b) = \{\{a\},\{a,b\}\}$. Esto garantiza —como el lector puede convencerse por sí mismo, probándolo— la igualdad de dos parejas según como se estableció arriba en el segundo inciso y, por ende, el significado de pareja ordenada; sin embargo, hemos decidido dejar el concepto en forma intuitiva en aras de simplicidad y brevedad.

Definición A.8 *Si A y B son dos conjuntos, se denota y define el producto cartesiano de estos conjuntos como*

$$A \times B = \{(a,b) \mid a \in A \text{ y } b \in B\}$$

▶ **Ejemplo A.10** Sean $A = \{1,2,3\}$, $B = \{a,b,c,d\}$, entonces

$$A \times B = \{(1,a),(1,b),(1,c),(1,d),(2,a),(2,b),(2,c),(2,d),(3,a),(3,b),(3,c),(3,d)\}. ◀$$

En general, $A \times B \neq B \times A$; como se puede ver en el ejemplo anterior. Por otra parte, los conceptos de parejas ordenadas y producto cartesiano se pueden extender de manera natural. En efecto, si α_i, β_i, $i = 1, \ldots, n$, son objetos, el símbolo $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ representa la n-eada ordenada con componentes los objetos α_i; y

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_n)$$

si y solamente si $\alpha_i = \beta_i$ para todo $i = 1, 2, \ldots, n$. Y si A_i son n conjuntos, se define y denota el producto cartesiano de ellos como

$$A_1 \times A_2 \times \cdots \times A_n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in A_i \text{ para cada } i = 1, 2, \ldots, n\}.$$

Leyes de De Morgan y propiedades algebraicas

Existen dos reglas bastante útiles y usadas con mucha frecuencia en teoría de conjuntos, llamadas leyes de De Morgan, que probamos en el siguiente teorema. Dejamos de ejercicio al lector la demostración de ellas mediante diagramas de Venn.

Teorema A.1 *Sean A y B dos conjuntos, entonces:*

1. $(A\cup B)^c = A^c\cap B^c$.
2. $(A\cap B)^c = A^c\cup B^c$.

DEMOSTRACIÓN ■ 1. Sea x cualquier elemento de $(A\cup B)^c$, entonces $x\notin A$ y $x\notin B$, por tanto $x\in A^c\cap B^c$; luego $(A\cup B)^c\subset A^c\cap B^c$. Sea $x\in A^c\cap B^c$ cualquier elemento, entonces $x\in A^c$ y $x\in B^c$, por tanto $x\notin A$ y $x\notin B$; luego $x\notin A\cup B$; es decir, $x\in (A\cup B)^c$ y, por ende, $A^c\cap B^c\subset (A\cup B)^c$. Hemos probado así que $(A\cup B)^c\subset A^c\cap B^c$ y $A^c\cap B^c\subset (A\cup B)^c$; por tanto $(A\cup B)^c = A^c\cap B^c$, de acuerdo con la definición A.2

2. $x\in (A\cap B)^c \Leftrightarrow x\notin A\cap B \Leftrightarrow x\notin A$ o $x\notin B \Leftrightarrow x\in A^c\cup B^c$; lo cual demuestra $(A\cap B)^c = A^c\cup B^c$. ■

Además de las leyes de De Morgan, las operaciones entre conjuntos tienen otras importantes propiedades; las principales se hacen patentes en la siguiente tabla:

De idempotencia	
1. $A\cup A = A$	2. $A\cap A = A$
De asociatividad	
3. $A\cup(B\cup C) = (A\cup B)\cup C$	4. $A\cap(B\cap C) = (A\cap B)\cap C$
De conmutatividad	
5. $A\cup B = B\cup A$	6. $A\cap B = B\cap A$
De distributividad	
7. $A\cup(B\cap C) = (A\cup B)\cap(A\cup C)$	8. $A\cap(B\cup C) = (A\cap B)\cup(A\cap C)$
De identidad	
9. $A\cup\emptyset = A$	10. $A\cap\mathscr{U} = A$
11. $A\cup\mathscr{U} = \mathscr{U}$	12. $A\cap\emptyset = \emptyset$
De complemento	
13. $A\cup A^c = \mathscr{U}$	14. $A\cap A^c = \emptyset$
15. $(A^c)^c = A$	16. $\mathscr{U}^c = \emptyset$, $\emptyset^c = \mathscr{U}$

Propiedades algebraicas de las operaciones de conjuntos

Finalmente, el siguiente teorema es bastante útil en la práctica y la demostración se deja como ejercicio al lector.

Teorema A.2 *Cada una de las siguientes condiciones son equivalentes a pares:*

1. $A\subset B$.
2. $A\cap B = A$.
3. $A\cup B = B$.
4. $B^c\subset A^c$.
5. $A\cap B^c = \emptyset$.

A.1.3 Reuniones e intersecciones de familias de conjuntos

Sea Λ un conjunto no vacío y supongamos que a cada λ corresponde un único conjunto A_λ; se forma entonces una familia de conjuntos $\{A_\lambda \mid \lambda \in \Lambda\}$. Una forma más compacta de representar esta familia es mediante la notación

$$\{A_\lambda\}_{\lambda\in\Lambda}$$

y se dice que ésta es una familia indexada de conjuntos y que Λ es un conjunto de índices para ella. Por ejemplo, si $\Lambda = \mathbb{N}$, el conjunto de números naturales, y para cada $n \in \mathbb{N}$ se define $A_n = [-1/n, 1/n]$, se obtiene la familia $\{A_n\}_{n\in\mathbb{N}}$. Sean ahora $\Lambda = \{\lambda \mid \lambda$ es un número telefónico$\}$ y $A_\lambda = \{x \mid x$ es dígito del número $\lambda\}$; luego, si $\lambda = 58645555$, entonces $A_\lambda = \{4,5,6,8\}$. Si $\{A_\lambda\}_{\lambda\in\Lambda}$ es una familia indexada de conjuntos, simplemente diremos que $\{A_\lambda\}_{\lambda\in\Lambda}$ es una familia de conjuntos.

Definición A.9 *Sea $\{A_\lambda\}_{\lambda\in\Lambda}$ una familia de conjuntos.*

1. *Se define y denota a reunión (unión) de esta familia como*

$$\bigcup_{\lambda\in\Lambda} A_\lambda = \{x \mid \textit{existe } \lambda\in\Lambda \textit{ tal que } x\in A_\lambda\}.$$

2. *Se define y denota la intersección de esta familia por*

$$\bigcap_{\lambda\in\Lambda} A_\lambda = \{x \mid x\in A_\lambda\ \forall\lambda\in\Lambda\}.$$

► **Ejemplo A.11** Se deja de ejercicio al lector probar lo siguiente:

1. Sea $A_n = [-1/n, 1/n]$, $n = 1, 2, \ldots$, entonces

$$\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} [-1/n, 1/n] = [-1, 1].$$

2. $$\bigcap_{n\in\mathbb{N}} A_n = \bigcap_{n\in\mathbb{N}} [-1/n, 1/n] = \{0\}.$$

3. $\Lambda = \{\lambda \mid \lambda$ es un número telefónico$\}$ y $A_\lambda = \{x \mid x$ es dígito del número $\lambda\}$, entonces

$$\bigcup_{\lambda\in\Lambda} A_\lambda = \{0,1,2,3,4,5,6,7,8,9\} \text{ y } \bigcap_{\lambda\in\Lambda} A_\lambda = \emptyset.$$ ◄

Cuando el conjunto de índices es $\Lambda = \mathbb{N}$, se acostumbra escribir $\bigcup_{n=1}^{\infty} A_n$ y $\bigcap_{n=1}^{\infty} A_n$ en lugar de $\bigcup_{n\in\mathbb{N}} A_n$ y $\bigcap_{n\in\mathbb{N}} A_n$, respectivamente. Notaciones similares se explican por sí solas para otros conjuntos numerables de subíndices, por ejemplo $\bigcup_{n=-\infty}^{\infty} A_n$ para la reunión $\bigcup_{n\in\mathbb{Z}} A_n$, o $\bigcap_{n=5}^{\infty} A_n$, etcétera.

Las leyes de De Morgan también son válidas para reuniones e intersecciones de familias de conjuntos:

$$\left(\bigcup_{\lambda\in\Lambda} A_\lambda\right)^c = \bigcap_{\lambda\in\Lambda} A_\lambda^c \quad \text{y} \quad \left(\bigcap_{\lambda\in\Lambda} A_\lambda\right)^c = \bigcup_{\lambda\in\Lambda} A_\lambda^c.$$

La demostración de estas dos relaciones es sencila y se deja de ejercicio al lector.

A.2 Demostraciones

Las matemáticas son un *juego* intelectual de muy alto nivel y, como todo juego, sólo se aprende *jugando*. En matemáticas —a diferencia de otros *juegos*, como la política— lo que se afirma se demuestra. Es imposible explicar en un conjunto de reglas cómo hacer demostraciones de proposiciones matemáticas. La mejor manera de aprender a realizar demostraciones matemáticas es ver cómo las hacen los matemáticos, en los libros de matemáticas, en casos específicos, y después intentar hacerlas por sí mismo, hasta que se puedan llevar a cabo demostraciones propias. Es decir, aprender a *jugar jugando*. Sin embargo, este oficio —el de probar proposiciones matemáticas— es especialmente delicado y por eso, en este segmento, se ha intentado bosquejar qué significa demostrar proposiciones matemáticas y cómo hacer estas demostraciones con casos específicos. Para ello primero primero esbozamos lo que es el método deductivo en matemáticas (en qué consiste el juego) y después explicamos algunas técnicas para demostrar teoremas; todo ello evitando teorizar —en la medida de lo posible— y tratando siempre casos concretos; pues el objetivo es que el lector tenga una guía para poder iniciarse en este apasionante y maravilloso (también tortuoso, ¿qué juego realmente interesante no lo es?) *juego* de probar proposiciones matemáticas.

A.2.1 El método deductivo

Las matemáticas son una ciencia deductiva; esto significa que parte de principios generales a hechos particulares. Los principios en los que se fundamenta una teoría matemática son llamados axiomas o postulados. Un axioma es una afirmación que se acepta como verdadera en una teoría, por su sencillez y carácter elemental. A partir de los axiomas se crea una cadena de resultados, llamados teoremas, que se deducen de los axiomas, teoremas que a su vez implican otros teoremas. Un teorema es una afirmación condicional que consiste de dos partes: una llamada hipótesis y otra tesis o conclusión. La hipótesis es una afirmación (o un conjunto de condiciones) y la tesis es otra afirmación (u otro conjunto de condiciones). La hipótesis y la tesis en un teorema se relacionan de la siguiente manera: si la hipótesis es verdadera (o el conjunto de condiciones de la hipótesis se cumple), entonces la tesis es verdadera (o el conjunto de condiciones de la tesis se cumple). Ésta es la razón por la que un teorema es una afirmación condicional. Un ejemplo de un axioma en geometría euclidiana es:

Axioma A.1 *Dados dos puntos distintos cualesquiera, hay exactamente una recta que los contiene.*

Mientras que un ejemplo de un teorema en geometría es:

Teorema A.3 *Si dos rectas diferentes se intersecan, entonces su intersección contiene solamente un punto.*

En este teorema la hipótesis es: *dos rectas diferentes se intersecan*; y la tesis: *su intersección contiene solamente un punto*.

En general un teorema tiene el formato:

$$\textit{Si } \mathbf{p}, \textit{entonces } \mathbf{q} \tag{A.1}$$

donde **p** es la hipótesis y **q** es la tesis. En un teorema la veracidad de la tesis se deduce de la veracidad de la hipótesis y es la razón del *si* condicional en (A.1). Es decir, **q** se deduce de **p**; lo cual se expresa diciendo que **p** implica **q** y se denota por

$$\mathbf{p} \Rightarrow \mathbf{q}.$$

Así, el teorema anterior se puede escribir como

dos rectas distintas se intersecan ⇒ *su intersección contiene solamente un punto*

En un teorema la tesis debe deducirse lógica y rigurosamente de la hipótesis y un teorema no es un axioma; es decir, la afirmación condicional que contiene no es verdadera *per se*; sino que debe demostrarse rigurosamente como consecuencia lógica de los axiomas y/o de otros teoremas que ya han sido probados antes. En una teoría matemática (por ejemplo geometría euclidiana) se dan definiciones para precisar conceptos y existen términos que no se definen por ser, como los axiomas, los más simples y fundamentales. Por ejemplo, rectas y puntos son términos que no se definen en geometría. Son entes elementales con los que se trabaja matemáticamente, pero que no es posible definirlos rigurosamente —trate, por ejemplo, de precisar lo que es punto y recta sin llegar a un círculo vicioso desde un punto de vista lógico, digamos: "un punto es la intersección de dos rectas" y "una recta es una sucesión de puntos que ..."—. Sin embargo, desde el enfoque matemático, es irrelevante qué son exactamente puntos y rectas, y lo que realmente importa es conocer qué relaciones geométricas guardan. Aunque desde el punto de vista matemático estos entes —puntos y rectas— no dejan de ser más que abstracciones de objetos del mundo físico como son, para un punto, la marca que deja en un papel un lápiz muy afilado o, para una recta, un hilo muy delgado y bien estirado, los teoremas de la geometría tienen una importante conexión con el mundo real y se han aplicado constantemente en diversas áreas,[1] regresando con ello al origen físico del que partieron; lo mismo sucede con otras ramas de las matemáticas. Por supuesto que los axiomas no tienen un carácter arbitrario —¿quién, en su sano juicio, los tomaría en serio si fuera así?— sino que describen ya sea una realidad física o una "realidad" matemática. Cuando se define un nuevo término en una teoría, se debe hacer utilizando términos que ya han sido definidos previamente. Sin embargo, la primera definición que se dé en una teoría no puede contener conceptos previamente definidos, pues no los hay; y, necesariamente, se tiene que hacer uso de los términos elementales indefinidos en la teoría. Lo mismo ocurre con el primer teorema, que, para ser probado, necesita sustentarse en teoremas previos que ya han sido demostrados pero que no existen aún en ese estado y, por ende, se tienen que utilizar

[1] Basta con observar la arquitectura para tener un ejemplo de la utilidad de la geometría.

los axiomas de la teoría que, a final de cuentas, también son teoremas pero que se aceptan como verdaderos sin demostración alguna. Por esta razón, los conceptos elementales no definidos y axiomas son imprescindibles en matemáticas. Un ejemplo de una definición primaria en geometría es:

Definición A.10 *Dos o más puntos son* colineales *si pertenecen a una misma recta.*

Aquí el concepto que se define es el de colinealidad y los términos no definidos en la teoría son punto y recta.

En general, si $\mathbf{p}$ es una afirmación y $\mathbf{q}$ es otra, podemos formar la proposición $\mathbf{p} \Rightarrow \mathbf{q}$ (nuevamente $\mathbf{p}$ es la hipótesis y $\mathbf{q}$ la tesis) que puede ser verdadera o falsa. Una teoría matemática, como mencionamos antes, se desarrolla a través de cadenas de proposiciones verdaderas $\mathbf{p}_1 \Rightarrow \mathbf{p}_2 \Rightarrow \cdots \Rightarrow \mathbf{p}_n \Rightarrow \cdots$, en forma deductiva. Algunas de estas proposiciones, por su trascendencia e importancia, son llamadas teoremas; las otras simplemente son parte de la cadena para concatenar toda la estructura deductiva y no son nombradas como teoremas. Para demostrar que una proposición es falsa, basta exhibir un contraejemplo; es decir, un caso particular que pruebe que no es cierta. Por ejemplo, la proposición: *Si A y B son conjuntos no vacíos, entonces $A \cap B$ es un conjunto no vacío* es falsa, pues los conjuntos $A = \{1\}$ y $B = \{2\}$ son no vacíos y $A \cap B = \emptyset$.

A.2.2 Métodos de demostración

Demostraciones directas

La demostración de una proposición $\mathbf{p} \Rightarrow \mathbf{q}$ es directa si existen proposiciones (verdaderas) $\mathbf{p}_1, \ldots, \mathbf{p}_k$ tales que $\mathbf{p} \Rightarrow \mathbf{p}_1 \Rightarrow \cdots \Rightarrow \mathbf{p}_k \Rightarrow \mathbf{q}$. A continuación damos unos ejemplos.

Sea $\mathbb{Z}$ el conjunto de números enteros; suponemos que se conocen las propiedades elementales de este conjunto como son las propiedades de las operaciones suma y producto.

Definición A.11 *Sean $a, b \in \mathbb{Z}$, con $a \neq 0$. Se dice que a divide al número b si existe $q \in \mathbb{Z}$ tal que $b = aq$. En tal caso se escribe $a \mid b$ y al entero a se le llama divisor de b. Si a no divide a b se denota este hecho por $a \nmid b$.*

Teorema A.4 *Sean $a, b, c \in \mathbb{Z}$, con $a \neq 0$. Si $a \mid b$ y $a \mid c$, entonces $a \mid mb + nc$ para todo par de enteros m y n.*

DEMOSTRACIÓN ■ Puesto que $a \mid b$ y $a \mid c$, entonces, por la definición A.11, existen un par de números enteros q_1 y q_2 tales que $b = aq_1$ y $c = aq_2$; por tanto, si $m, n \in \mathbb{Z}$ son cualquier par de números,

$$\begin{aligned} mb + nc &= m(aq_1) + n(aq_2) \\ &= a(mq_1 + nq_2) \end{aligned}$$

y ya que $mq_1 + nq_2 \in \mathbb{Z}$ porque $q_1, q_2, m, n \in \mathbb{Z}$, se deduce que $a \mid mb + nc$. ■

Otra forma de escribir esta demostración es: $a \mid b$ y $a \mid c \Rightarrow b = dq_1$ y $c = dq_2$ para ciertos $q_1, q_2 \in \mathbb{Z} \Rightarrow mb + nc = m(aq_1) + n(aq_2) = aq$ con $q = (mq_1 + nq_2) \in \mathbb{Z} \Rightarrow a \mid mb + nc$. Por lo que es una demostración directa. Observe que cada paso de la demostración se argumentó rigurosamente utilizando la definición de divisibilidad A.11, el hecho de que la multiplicación y la suma de números enteros es cerrada en este conjunto y las propiedades de las operaciones en los números enteros.

Para dar un ejemplo más, aceptaremos también como demostrado el conocido algoritmo de división: *si* $a, b \in \mathbb{Z}$, *con* $a \neq 0$, *entonces existen únicos enteros* q *y* r *tales que* $b = aq + r$ *y* $0 \leq r < |a|$; y utilizaremos el siguiente axioma de los números naturales:

Axioma A.2 (Principio del buen orden) *Sea* B *un subconjunto no vacío de los números naturales* $\mathbb{N} = \{1, 2, 3, \ldots\}$, *entonces existe un elemento mínimo de* B; *es decir, existe* $m_0 \in B$ *tal que* $m_0 \leq b$ *para todo* $b \in B$.

Observe que este axioma es simple, muy elemental y *evidente*; que son las características de un axioma. También conviene notar que la siguiente definición contiene términos previamente definidos (el concepto de divisibilidad de la definición A.11).

Definición A.12 *Sean* $a, b, d \in \mathbb{Z}$, *con* $d > 0$. *Se dice que* d *es un máximo común divisor de* a, b *si:*

1. $d \mid a$ *y* $d \mid b$
2. *Si* d' *es un entero que divide a* a *y* b, *entonces* $d' \mid d$.

Notemos que si a, b tienen un máximo común divisor, entonces éste es único; pues en tal caso si d_1 es otro máximo común divisor, entonces $d' \mid d$ y $d \mid d'$, lo cual implica[2] $d = d'$. Ahora, en el siguiente teorema, probaremos que si a, b son dos enteros no ambos iguales a cero, entonces existe el máximo común divisor de ellos y se puede escribir como una combinación lineal de a y b.

Teorema A.5 *Sean* $a, b \in \mathbb{Z}$ *no ambos cero. Entonces existen* $m_0, n_0 \in \mathbb{N}$, *tales* $d = m_0 a + n_0 b$ *es el máximo común divisior de* a *y* b.

DEMOSTRACIÓN ■ Sea $\mathfrak{M} = \{ma + nb : n, m \in \mathbb{Z}\}$. Notemos que $a, b \in \mathfrak{M}$ y que si $u = ma + nb \in \mathfrak{M}$, entonces $-u \in \mathfrak{M}$. Ya que uno de los enteros a y b es distinto de 0, se desprende que $\mathfrak{M}$ contiene enteros positivos. Denotemos por B a este conjunto (no vacío); por el principio del buen orden (axioma A.2), existe un elemento mínimo $d = m_0 a + n_0 b \in B$. Por el teorema A.4 si $d' \mid a$ y $d' \mid b$, entonces $d' \mid m_0 a + n_0 b = d$. Lo cual prueba la condición 2 de la definición A.12. Por el algoritmo de división existen $q, r \in \mathbb{Z}$ tales que

$$a = dq + r$$

con $0 \leq r < d$. Luego,

$$a = (m_0 a + n_0 b)q + r$$

[2] Como uno divide al otro, existen dos enteros q_1 y q_2 tales que $d = d'q_1$ y $d' = dq_2$; por tanto $d = d(q_1 q_2)$ y, entonces, $d(1 - q_1 q_2) = 0$; puesto que d y d' son positivos (por la definición de máximos comunes divisores de a y b), q_1 y q_2 son positivos. Entonces, de la igualdad $d(1 - q_1 q_2) = 0$, se desprende $q_1 q_2 = 1$ y, por ende, $q_1 = 1 = q_2$; luego, $d = d'$.

de donde

$$r = a(1 - m_0 q) + (-n_0 q)b$$

Ya que el menor elemento de B es d, la posibilidad $r > 0$ no se puede dar pues $r < d$, por tanto se tiene $r = 0$ y, entonces, $d \mid a$. Intercambiando los papeles de a y b en lo precedente, se demuestra que $d \mid b$. Luego $d = m_0 a + n_0 b$ también cumple con la condición 1 de la definición A.12 y en consecuencia es el máximo común divisor de a y b. ■

En lo precedente dimos definiciones, teoremas y términos indefinidos; pues el propio concepto de número entero está indefinido. Sin embargo, los números enteros se pueden definir utilizando conceptos mucho más fundamentales, que, como mencionamos antes, requerirán, aun así, de conceptos más elementales no definidos.

Demostraciones indirectas

Probemos el teorema A.3: Sean L_1 y L_2 dos rectas, que por hipótesis hemos supuesto distintas. Ahora supongamos que la tesis es falsa; es decir que la intersección de L_1 y L_2 contiene por lo menos dos puntos distintos, digamos que p_1 y p_2 son un par de ellos. Entonces L_1 y L_2 son dos rectas que pasan por los puntos distintos p_1 y p_2; luego, por el axioma A.1, L_1 y L_2 deben ser la misma recta; lo que nos lleva a probar que la hipótesis $L_1 \neq L_2$ es falsa. Por tanto p_1 y p_2 no pueden ser distintos y, entonces, son iguales.

A este tipo de demostración se le llama indirecta. En el caso general, para utilizar esta técnica y probar una proposición $\mathbf{p} \Rightarrow \mathbf{q}$, se demuestra la proposición $\sim \mathbf{q} \Rightarrow \sim \mathbf{p}$, si ésta resulta verdadera, entonces la proposición $\mathbf{p} \Rightarrow \mathbf{q}$ también (y viceversa), donde $\sim \mathbf{r}$ es la negación de la afirmación $\mathbf{r}$. Para profundizar un poco más en este tema mostremos primero la siguiente afirmación.

Proposición A.1 *Si $a \in \mathbb{Z}$ un número dado y p es un número primo*[3] *que no divide a a, entonces p y a son primos relativos; es decir, su máximo común divisor es* 1.

DEMOSTRACIÓN ■ Sea $d \geq 1$ un divisor de a y p, entonces, como $d \mid p$, se tiene $d = 1$ o $d = p$, por ser p un número primo. Pero ya que por hipótesis $p \nmid a$, d no puede ser p y, por tanto, $d = 1$; luego a y p son primos relativos. ■

Sean $P = \{x : x \text{ es un número primo y } x \nmid a\}$ y $Q = \{m : a \text{ y } m \text{ son primos relativos}\}$. Sea la afirmación $\mathbf{s}$: *p es un número primo y $p \nmid a$* ; y sea la afirmación $\mathbf{q}$: *p y a son primos relativos*. Entonces claramente la proposición $\mathbf{s} \Rightarrow \mathbf{q}$ es verdadera si $P \subset Q$. Como la implicación $\mathbf{s} \Rightarrow \mathbf{q}$ es verdadera, podemos expresar esta proposición diciendo que una condición necesaria para que $p \in P$ es que $p \in Q$, o, equivalentemente: una condición necesaria para que un número primo no divida a un número a es que sean primos relativos. Es decir, una condición necesaria para que $\mathbf{s}$ sea verdadera es que $\mathbf{q}$ sea verdadera. O también se puede redactar como: $\mathbf{s}$ sólo si $\mathbf{q}$.

En general, en una proposición de la forma $\mathbf{p} \Rightarrow \mathbf{q}$, podemos formar los conjuntos P y Q cuyos elementos, respectivamente, son aquellos para los cuales $\mathbf{p}$ es verdadera y $\mathbf{q}$ es verdadera. Entonces la

[3] Cfr. definición A.13, página 1000.

veracidad de la proposición $\mathbf{p} \Rightarrow \mathbf{q}$ equivale a $P \subset Q$. Y puesto que $P \subset Q$ equivale (cfr. teorema A.2, pág. 991) a $Q^c \subset P^c$, la proposición $\mathbf{p} \Rightarrow \mathbf{q}$ equivale a la proposición $\sim \mathbf{q} \Rightarrow \sim \mathbf{p}$; donde, como antes, $\sim \mathbf{s}$ significa la negación de la proposición **s**. Y se tienen, por el mismo argumento dado arriba, las formas equivalentes de plantear la proposición $\mathbf{p} \Rightarrow \mathbf{q}$: *(a)* **p** sólo si **q**; *(b)* **q** es necesario para **p**; *(c)* **q** si **p**.

Demostraciones por reducción al absurdo

La demostración del teorema A.3 se puede enfocar desde otra perspectiva equivalente. Siguiendo el hilo del razonamiento dado arriba para probar este teorema, se concluye que existen dos rectas distintas (pues la hipótesis es verdadera) que pasan por dos puntos diferentes p_1 y p_2, lo cual es una contradicción al axioma A.1 que es verdadero; o una contradicción a la hipótesis, si aplicamos dicho axioma, que se supone es verdadera: un *absurdo*. Entonces, para aplicar esta técnica se niega la tesis y se llega a la contradicción de la hipótesis, o de algún axioma o teorema ya demostrado; entonces, el suponer falsa la tesis lleva a un hecho contradictorio, un absurdo, y, por tanto, la tesis no pueder ser falsa, luego es verdadera.

Veamos un ejemplo más, la irracionalidad del número $\sqrt{2}$. Antes necesitamos establecer unos resultados preliminares. Recordemos que un número entero n es par si $n = 2k$ para algún k entero y es impar si $n = 2k_1 + 1$ para cierto entero k_1.

Lema A.1 *Sea n un número entero. Entonces:*

1. *n es par $\Rightarrow n^2$ es par.*
2. *n es impar $\Rightarrow n^2$ es impar.*
3. *n^2 es par $\Rightarrow n$ es par.*
4. *n^2 es impar $\Rightarrow n$ es impar.*

DEMOSTRACIÓN ■ 1. Si n es par, entonces existe $k \in \mathbb{Z}$ tal que $n = 2k$, luego $n^2 = (2k)^2 = 4k^2 = 2(2k^2)$, y como $2k^2 \in \mathbb{Z}$ porque $2, k \in \mathbb{Z}$, se desprende que n^2 es par.

2. Si n es impar existe $k \in \mathbb{Z}$ tal que $n = 2k+1$, luego $n^2 = (2k+1)^2 = 4k^2 + 2k + 1 = 2(2k^2 + k) + 1$; y ya que $2k^2 + k \in \mathbb{Z}$ porque $2, k \in \mathbb{Z}$, se tiene que n^2 es impar.
3. Si n^2 es par,[4] n debe ser par, pues en caso contrario, por el inciso 2, n^2 sería impar; lo cual contradice la hipótesis.
4. Si n^2 es impar, n debe ser impar, pues en caso contrario, por el inciso 1, n^2 sería par; lo cual contradice la hipótesis. ■

Recordemos que $\mathbb{Q}$ es el conjunto de los números racionales; esto es, $\mathbb{Q} = \{p/q : p, q \in \mathbb{Z} \text{ y } q \neq 0\}$.

Teorema A.6 $\sqrt{2} \notin \mathbb{Q}$.

DEMOSTRACIÓN ■ Procedamos por reducción al absurdo. Supongamos que $\sqrt{2} \in \mathbb{Q}$, entonces existen $p, q \in \mathbb{Z}$, con $q \neq 0$, tales que

[4] Aceptaremos como cierta la proposición de que todo número entero es un número par o un número impar; la cual el lector puede probar fácilmente usando el principio de inducción del apartado A.3.

$$\sqrt{2} = \frac{p}{q}.$$

Podemos suponer que p y q son primos relativos; es decir, el máximo común divisor de ellos es[5] 1. Al elevar al cuadrado ambos lados de la igualdad precedente se obtiene

$$\frac{p^2}{q^2} = 2$$

y, por tanto,

$$p^2 = 2q^2.$$

Entonces p^2 es par y, por el lema anterior, p también es par. Luego $p = 2k$ para algún $k \in \mathbb{Z}$, con lo que

$$\begin{aligned} 2(2k^2) &= 4k^2 \\ &= 2q^2 \end{aligned}$$

y, por ende,

$$q^2 = 2k^2$$

es decir, q^2 es un número par y, en consecuencia, por el lema anterior, q también es par. Entonces p y q tienen a 2 por común divisor, lo cual es una contradicción a la suposición de que p y q son primos relativos. Por tanto $\sqrt{2} \notin \mathbb{Q}$. ■

Observe que el teorema precedente no tiene el formato (A.1); sin embargo, se puede escribir como: *Si* $p/q \in \mathbb{Q}$, *entonces* $p/q \neq \sqrt{2}$.

A.2.3 Bicondicional y definiciones, lemas y corolarios

Toda proposición $\mathbf{p} \Rightarrow \mathbf{q}$ tiene una proposición recíproca, que es la que se obtiene al invertir la hipótesis y la tesis; esto es, $\mathbf{q} \Rightarrow \mathbf{p}$. Cuando ambas son verdaderas, se escribe $\mathbf{p} \Leftrightarrow \mathbf{q}$, que también se representa como **p** *si y sólo si* **q**. En tal caso se dice que las afirmaciones **p** y **q** son equivalentes. Por ejemplo:

> *Un triángulo es rectángulo* si y sólo si *la suma de los cuadrados de sus catetos es el cuadrado de su hipotenusa.*

Se tiene, para la proposición $\mathbf{p} \Leftrightarrow \mathbf{q}$:

1. A la proposición $\mathbf{p} \Leftrightarrow \mathbf{q}$ se le llama bicondicional porque equivale a $\mathbf{p} \Rightarrow \mathbf{q}$ y $\mathbf{q} \Rightarrow \mathbf{p}$.
2. Como en la proposición (verdadera) condicional $\mathbf{p} \Rightarrow \mathbf{q}$, la afirmación **q** es verdadera si la afirmación **p** lo es; y **p** es verdadera solamente si la afirmación **q** es verdadera, la proposición bicondicional se escribe: **p** *si y sólo si* **q**.

[5] Si no son primos relativos dividimos el numerador y el denominador por d, el máximo común divisor de ellos, obteniendo el mismo número pero con numerador y denominador números primos relativos.

3. Puesto que en la proposición condicional $\mathbf{p} \Rightarrow \mathbf{q}$, una condición necesaria para que **p** sea verdadera es que **q** sea verdadera; y en la implicación válida $\mathbf{q} \Rightarrow \mathbf{p}$, la veracidad de **q** también es suficiente para la veracidad de **p**; la proposición bicondicional $\mathbf{p} \Leftrightarrow \mathbf{q}$ se acostumbra también expresar como: **q** *es una condición necesaria y suficiente para* **p** (y viceversa).

Toda definición matemática debe entenderse en sentido bicondicional; consideremos, por ejemplo, la siguiente definición:

Definición A.13 *Un número entero $p > 1$ es un número **primo** si los únicos divisores positivos de él son 1 y p.*

Significa que si $q > 1$ tiene por divisores positivos únicamente a 1 y a q, entonces q se llama número primo; inversamente si de alguna forma se sabe que un número p es primo, entonces él es mayor a uno, y los únicos divisores de p son la unidad y él mismo. La razón de hacer énfasis en esta idea (que probablemente resulte obvia al lector) es que existen formas equivalentes de describir un mismo concepto. El siguiente teorema sirve para ilustrar este hecho y, de paso, es un ejemplo de una proposición bicondicional que probaremos.

Teorema A.7 *Para que p sea un número primo es necesario y suficiente que se cumpla la siguiente condición: $p > 1$ y si $a, b \in \mathbb{Z}$ y $p \mid ab$ entonces $p \mid a$ o $p \mid b$.*

DEMOSTRACIÓN ■ ($\Rightarrow$) Necesidad: Supongamos que p es un número primo. Sean $a, b \in \mathbb{Z}$ y supongamos que $p \mid ab$, entonces existe $q \in \mathbb{Z}$ tal que $ab = qp$. Si $p \nmid a$, por la proposición A.1, p y a son primos relativos; así, por el teorema A.5, existen $m_0, n_0 \in \mathbb{Z}$ tales que

$$1 = m_0 a + n_0 p$$

y, por tanto,

$$\begin{aligned} b &= m_0 ab + n_0 pb \\ &= m_0(qp) + n_0 pb \\ &= (m_0 q + n_0 b)p \end{aligned}$$

lo cual implica (pues $m_0 q + n_0 b \in \mathbb{Z}$) que p divide a b.
($\Leftarrow$) Suficiencia: Ahora supongamos que $p > 1$ y que para todo $a, b \in \mathbb{Z}$, $p \mid ab \Rightarrow p \mid a$ o $p \mid b$. Sea $d \geq 1$ un entero que es divisor de p. Si p divide a d, entonces (cfr. nota al pie de la página 996) $p = d$; si p no divide a d, por la proposición A.1, p y d son primos relativos; y ya que d divide a p, se desprende que $d \mid 1$ y, puesto que d es positivo, se concluye que $d = 1$. Se ha probado así que cualquier divisor positivo de p es la unidad o el propio $p > 1$, luego p es un número primo. ■

Una proposición puede tener más de dos afirmaciones equivalentes. Diremos que las condiciones $\mathbf{q}_i$, $i = 1, 2, \ldots, n$, son equivalentes a pares si cuando se cumple una, cualquiera de ellas, entonces se cumplen todas las demás; esto es, $\mathbf{q}_i \Leftrightarrow \mathbf{q}_j$ para todo $i, j \in \{1, 2, \ldots, n\}$, o $\mathbf{q}_1 \Rightarrow \mathbf{q}_2 \Rightarrow \cdots \Rightarrow \mathbf{q}_n \Rightarrow \mathbf{q}_1$. Por ejemplo:

Teorema A.8 *Sea $p > 1$ un entero positivo. Las siguientes afirmaciones son equivalentes a pares:*

1. *p es un número primo.*
2. *$p \mid ab \Rightarrow p \mid a$ o $p \mid b$ para todo $a, b \in \mathbb{Z}$.*
3. *$p \nmid a \Rightarrow p$ y a son primos relativos; para todo $a \in \mathbb{Z}$.*

De hecho ya hemos probado la equivalencia de las condiciones 1 y 2; se deja al lector de ejercicio terminar la demostración de este teorema de preferencia con sus propias palabras; por ejemplo, probando (1)⇒(2)⇒(3)⇒(1).

Es evidente que el recíproco de un teorema no necesariamente es cierto; es decir, puede ser que $\mathbf{p} \Rightarrow \mathbf{q}$, pero $\mathbf{q} \not\Rightarrow \mathbf{p}$; como hacemos patente en el siguiente teorema y contraejemplo.

Antes necesitamos de los dos siguientes lemas, los cuales el lector puede probar; el primero siguiendo exactamente el argumento de demostración de la proposición A.7; y en el segundo puede utilizar el teorema A.4 y el conocido teorema del binomio de Newton[6] o ver la demostración que se da mediante el principio de inducción en el siguiente apartado.[7]

Lema A.2 *Sean $a, b, c \in \mathbb{Z}$; con $a \neq 0$, si a y b son primos relativos y $a \mid bc$, entonces $a \mid c$.*

Lema A.3 *Si a y b son dos enteros primos relativos, entonces a^m y b son primos relativos para todo $m = 1, 2, \ldots$*

Teorema A.9 *Sea $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ un polinomio de coeficientes enteros y de grado $n \geq 1$. Si $\frac{p}{q} \in \mathbb{Q}$ es una raíz racional de este polinomio; esto es, $f(p/q) = 0$, entonces $p \mid a_0$ y $q \mid a_n$.*

DEMOSTRACIÓN ■ Podemos suponer, sin perder generalidad, que p y q son primos relativos (cfr. nota al pie de la página 999). Entonces, por hipótesis,

$$
\begin{aligned}
0 &= f(p/q) \\
&= a_n \frac{p^n}{q^n} + a_{n-1}\frac{p^{n-1}}{q^{n-1}} + \cdots + a_1 \frac{p}{q} + a_0.
\end{aligned}
$$

Al multiplicar por q^n ambos lados de la precedente igualdad se obtiene

$$a_n p^n + a_{n-1} q p^{n-1} + \cdots + a_1 q^{n-1} p + a_0 q^n = 0 \tag{A.2}$$

de donde, al ser $n \geq 1$,

$$a_0 q^n = p(-a_n p^{n-1} - a_{n-1} q p^{n-2} - \cdots - a_1 q^{n-1})$$

[6] Cfr. ejemplo A.15, página 1006.

[7] Cfr. ejemplo A.14 en la página 1005.

y, por tanto, $p \mid a_0q^n$. Ya que p y q son primos relativos, por el lema A.3, p y q^n son primos relativos; y puesto que $p \mid a_0q^n$, se desprende, del lema A.2, que $p \mid a_0$. Al despejar a_np^n de la relación (A.2) resulta

$$a_np^n = q(-a_{n-1}p^{n-1} - \cdots - a_1q^{n-2}p - a_0q^{n-1})$$

de donde $q \mid a_np^n$; como p y q son primos relativos los lemas A.3 y A.2 implican $q \mid a_n$. ■

Evidentemente el recíproco del teorema anterior no es cierto. Por ejemplo, para el polinomio $f(x) = x^2+1$, $a_2 = 1$ y $a_0 = 1$; así, los números racionales p/q con p divisor de $a_0 = 1$ y q divisor de $a_2 = 1$, son $p/q = \pm 1$; pero ninguno de estos valores es raíz del polinomio $f(x)$. Es decir, la condición $p \mid a_0$ y $q \mid a_n$ es necesaria pero no suficiente para que el número racional p/q sea raíz del polinomio con coeficientes enteros $f(x)$.

Lemas y corolarios

Un lema es un teorema que se plantea *ex profeso* para ayudar a probar uno o varios teoremas. Así, por ejemplo, los lemas A.2 y A.3 se establecieron para probar el teorema A.9. Un corolario es un teorema que se deduce inmediatamente de un teorema previo. Por ejemplo:

Corolario A.1 (Corolario del teorema A.5) *$a, b \in \mathbb{Z}$ son primos relativos si y sólo si existen $n_0, m_0 \in \mathbb{Z}$ tales que $n_0a + m_0b = 1$.*

DEMOSTRACIÓN ■ ($\Rightarrow$) Si $a, b \in \mathbb{Z}$ son primos relativos, entonces 1 es su máximo común divisor y, por el teorema A.5, existen $n_0, m_0 \in \mathbb{Z}$ tales que $n_0a + m_0b = 1$. ($\Leftarrow$) Supongamos que existen dos enteros, n_0 y m_0, tales que $n_0a + m_0b = 1$. Entonces, si d es un entero positivo divisor de a y b, por el teorema A.4, $d \mid n_0a + m_0b = 1$; luego, $d \mid 1$ y, por tanto, $d = 1$; es decir a y b son primos relativos. ■

Naturalmente el uso del término lema o corolario es relativo, y depende con frecuencia del gusto del autor; porque en muchas ocasiones un lema en realidad tiene tanta o más importancia que los teoremas cuya demostración sustenta; sin embargo, por ese mismo hecho se resalta la denominación de lema, para tenerlo en mente a lo largo de la estructura matemática subyacente y diferenciarlo de los demás teoremas. Lo mismo puede suceder con un corolario y habrá quien lo llame así o prefiera establecerlo como teorema.

A.3 Inducción matemática

El proceso de inducción es una forma de razonar que parte de hechos particulares a principios generales. La mayoría de las ciencias (exceptuando las matemáticas) utiliza este proceso cognitivo para producir leyes generales con base en observaciones particulares. Por ejemplo, si se realizan varios experimentos en los cuales se arrojan diversos tipos de cuerpos desde una misma altura cerca de la superficie terrestre y se observa que todos caen, al momento de tocar tierra, con la misma velocidad, entonces se generaliza este hecho a cualquier cuerpo diciendo: *todos los cuerpos que son arrojados desde una misma altura (no muy lejana de la superficie terrestre) caerán a la tierra con la misma velocidad independientemente de su forma y tamaño*. Mientras todos los experimentos que se realicen para comprobar este hecho produz-

can el mismo resultado, esta “ley” será cierta; a menos que un experimento particular, en condiciones equivalentes, dé un resultado distinto. En matemáticas no se puede hacer lo mismo porque ni siquiera es una ciencia experimental. Aunque basa, como toda ciencia, su desarrollo en la experiencia humana, el que un hecho valga para muchos casos y no se encuentre uno en el que no sea cierto, no se puede tomar como válido para todos los demás; sino que tiene que ser probado rigurosamente por medio del método deductivo. El siguiente es un ejemplo del porqué no se pueden hacer generalizaciones en matemáticas de evidencias particulares sin hacer uso del método deductivo.

► **Ejemplo A.12** Recordemos que un número entero $q > 1$ es un número primo si los únicos divisores positivos de p son 1 y el propio p. Consideremos la fórmula $p(n) = n^2 - n + 41$. Entonces $p(1) = 41$, $p(2) = 43$, $p(3) = 47$, $p(4) = 53$, $p(5) = 61$; los cuales son todos números primos. Podríamos conjeturar ingenuamente que $p(n)$ es un número primo para todo $n \in \mathbb{N}$. De hecho el lector puede comprobar por sí mismo que $p(n)$ es un número primo para todo entero n con $1 \leq n \leq 40$; sin embargo, $p(41) = 41^2$ ya no es un número primo.◄

Aunque el caso precedente es muy discreto —es fácil observar de inmediato que $p(41)$ no es un número primo—, sirve para ilustrar la idea general: puede existir una afirmación $p(n)$ que sea verdadera para todos los valores enteros desde 1 hasta valores muy grandes de n; pero que para un valor mayor, n_0, $p(n_0)$ ya no sea verdadera y que, por limitaciones humanas, no sea posible llevar a efecto los cálculos hasta ese número, quizá desmesuradamente grande. Un ejemplo famoso de esta situación es la conjetura de Goldbach; en ella se afirma que todo número entero par mayor que 2 es la suma de dos números primos: $4 = 2 + 2$, $6 = 3 + 3$; $8 = 3 + 5$, $10 = 5 + 5$, $12 = 5 + 7$, $14 = 7 + 7$, $16 = 3 + 13$, $18 = 5 + 13$, etc. Desde 1742 —cuando C. Goldbach dio a conocer esta conjetura a L. Euler— hasta el 2008, no se ha encontrado un número par mayor a 2 y menor que 12×10^{17}que no cumpla con la conjetura y la lista sigue aumentando a la fecha.

Sin embargo, supongamos ahora que tenemos una afirmación $p(n)$, $n \in \mathbb{N}$, que satisface estas dos condiciones: *(i)* $p(1)$ es verdadera y, para cualquier $n \in \mathbb{N}$, *(ii)* si $p(n)$ es verdadera, $p(n+1)$ también es verdadera. Entonces tendríamos lo siguiente: como $p(1)$ es verdadera por *(i)*, $p(2)$ es verdadera por *(ii)*; ya que $p(2)$ es verdadera, por *(ii)*, $p(3)$ es verdadera; y así *ad infinítum*; esto es, $p(n)$ es verdadera para todo $n \in \mathbb{N}$. Por supuesto que ésta no es de ninguna manera una argumentación rigurosa; pero se puede establecer, ya que parece ser bastante evidente y simple, como un postulado, el llamado *principio de inducción matemática* que enunciamos a continuación:

Principio de inducción: *Sea $p(n)$ una afirmación cuyo valor de verdad (falso o verdadero) depende del número natural n. Supongamos que se cumplen las siguientes dos condiciones:*

(i) *$p(1)$ es verdadera.*
(ii) *Para cualquier $k \in \mathbb{N}$, suponer que $p(k)$ es verdadera implica que $p(k+1)$ es verdadera.*

Entonces $p(n)$ es verdadera para todo $n \in \mathbb{N}$.

A la suposición: $p(k)$ *es verdadera*, se le llama hipótesis de inducción cuando se usa este principio.

► **Ejemplo A.13** De niño el matemático C. F. Gauss hizo la siguiente conjetura: para todo $n \in \mathbb{N}$, se tiene

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2} \qquad \text{(A.3)}$$

La conjetura se basa en el hecho de que al sumar, por ejemplo, los primeros 100 enteros naturales,

$$1+2+\cdots+98+99+100$$

podemos agrupar de la siguiente manera:

$$\begin{aligned} 1+2+\cdots+50+51+98+99+100 &= (1+100)+(2+99)+(3+98)+\cdots+(50+51) \\ &= 50(101) \\ &= \frac{100}{2}(100+1). \end{aligned}$$

Ahora probemos la fórmula (A.3) utilizando el principio de inducción. En este caso $p(n)$ es la afirmación: *la fórmula* (A.3) *es verdadera para* n. Entonces, ya que

$$1 = \frac{1(1+1)}{2},$$

$p(1)$ es verdadera. Sea $k \in \mathbb{N}$ y supongamos que $p(k)$ es verdadera para este valor; esto es, (A.3) es válida para $n = k$. Luego, por hipótesis de inducción,

$$\begin{aligned} 1+2+\cdots+k+(k+1) &= (1+2+\cdots+k)+(k+1) \\ &= \frac{k(k+1)}{2}+(k+1) \\ &= \frac{k(k+1)+2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2} \\ &= \frac{(k+1)((k+1)+1)}{2}. \end{aligned}$$

Así, el suponer que $p(k)$ es verdadera implica que $p(k+1)$ es verdadera; por el principio de inducción (A.3) es válida para todo $n \in \mathbb{N}$.◀

Podemos escribir en un lenguaje menos vago el principio de inducción utilizando el lenguaje de la teoría de conjuntos, de la siguiente manera:

Principio de inducción (con notación de conjuntos): *Sea A un conjunto de números naturales; es decir, $A \subset \mathbb{N}$, que satisface las siguientes dos condiciones:*

(i) $1 \in A$.
(ii) $n \in A \Rightarrow n+1 \in A$.

Entonces $A = \mathbb{N}$.

Para ello basta que A sea el conjunto de números enteros $n \geq 1$ para los que $p(n)$ es verdadera.

Aunque hemos establecido el principio de inducción como un axioma, en realidad lo podemos deducir del principio del buen orden (axioma A.2) y establecerlo como un teorema. Pero si se elige el principio de inducción como axioma, entonces se puede deducir el principio del buen orden a partir del principio de inducción, como un teorema. En el teorema A.10 hacemos patente la equivalencia de estos dos

principios. Entonces la elección de cuál fijar como axioma y cuál como teorema es cuestión de gustos. En lo personal, prefiero fijar como axioma el principio del buen orden por su enorme simplicidad, porque tiene una formulación mucho más sencilla y es *evidente*; de hecho, se escribe en una sola línea: *todo subconjunto no vacío de números naturales tiene un elemento mínimo*. Obviamente el lector que así lo desee puede fijar ambos como axiomas, omitir la lectura del siguiente teorema y continuar con los ejemplos de abajo; mas se debe recordar que en este apéndice el objetivo es aprendrer a realizar demostraciones matemáticas y mientras más ejemplos se tengan a la mano, será mejor.

Teorema A.10 *Las siguientes condiciones son equivalentes a pares:*

1. *El principio del buen orden*
2. *El principio de inducción*

DEMOSTRACIÓN ■ ((1)⇒(2)) Supongamos que se cumple el principio del buen orden. Sea $A \subset \mathbb{N}$ un conjunto que satisface *(i)* $1 \in A$ y *(ii)* $n \in A \Rightarrow n+1 \in A$. Sea $B = A^c$, el complemento de A relativo al conjunto $\mathbb{N}$. Supongamos que $B \neq \emptyset$, entonces, por el principio del buen orden, B tiene un elemento mínimo n_0 que es mayor a 1 porque $1 \in A$ debido a la condición *(i)*; luego $n_0 - 1 \in \mathbb{N}$ y, entonces, $n_0 - 1 \notin B$, pues $n_0 - 1 < n_0$ y n_0 es el mínimo en B. Por tanto $n_0 = (n_0 - 1) + 1 \in A$, por la condición *(ii)*. De esta manera, $n_0 \in A \cap B = A \cap A^c = \emptyset$; lo cual es una contradicción. Entonces B debe ser el conjunto vacío y en consecuencia $A = B^c = \mathbb{N}$.
((2)⇒(1)) Supongamos ahora que se cumple el principio de inducción. Sea $A \subset \mathbb{N}$ y supongamos que A no tiene un elemento mínimo. Sea $B = \{n \in \mathbb{N} \mid 1, 2, \ldots, n \notin A\}$. Entonces $1 \in B$, porque A no tiene elemento mínimo. Sea $n \in B$, entonces $1, 2, \ldots, n \notin A$, por tanto $n+1 \notin A$, pues en caso contrario $n+1$ sería el mínimo de A. Luego el conjunto B cumple: *(i)* $1 \in B$ y *(ii)* $n \in B \Rightarrow n+1 \in B$; así, por el principio de inducción, $B = \mathbb{N}$; lo cual implica $A = \emptyset$. Entonces todo conjunto no vacío de los números naturales debe tener un elemento mínimo. ■

► **Ejemplo A.14** Demostrar el lema A.3; es decir, probar que si $a, b \in \mathbb{Z}$ son primos relativos, entonces a^m y b son también primos relativos para todo $m \in \mathbb{N}$.◄

DEMOSTRACIÓN ■ El lema es cierto para $m = 1$ por la hipótesis de que a y b son primos relativos. Sea $m \in \mathbb{N}$ y supongamos que a^m y b son primos relativos (hipótesis de inducción). Entonces, por la hipótesis de inducción y el hecho de que a y b son primos relativos, el teorema A.5 (p. 996) implica que existen $n_0, m_0, n_1, m_1 \in \mathbb{Z}$ tales que

$$\begin{aligned} 1 &= n_0 a + m_0 b \quad \text{y} \\ 1 &= n_1 a^m + m_1 b \end{aligned}$$

Luego,

$$\begin{aligned} 1 &= (n_0 a + m_0 b)(n_1 a^m + m_1 b) \\ &= (n_0 n_1) a^{m+1} + (n_0 m_1 a + m_0 n_1 a^m + m_0 b m_1) b \end{aligned}$$

y, por el corolario A.1 (p. 1006), se desprende que a^{m+1} y b son primos relativos. Así, por el principio de inducción, a^n y b son primos relativos para todo $n \in \mathbb{N}$. ■

Sea $n_0 \in \mathbb{Z}$, denotaremos al conjunto de todos los enteros a partir de n_0 con el símbolo $[\![n_0, \infty[\![$, el intervalo de todos los números enteros a partir de n_0; esto es

$$[\![n_0, \infty[\![= \{m \in \mathbb{Z} \mid m \geq n_0\}.$$

De manera análoga, si $n, m \in \mathbb{Z}$, $m < n$, se define el intervalo cerrado de números enteros

$$[\![m, n]\!] = \{x \in \mathbb{Z} \mid m \leq x \leq n\}.$$

El principio de inducción puede generalizarse si en lugar de los números naturales se toma el conjunto $[\![n_0, \infty[\![$.

Principio de inducción (con corrimiento): *Sea $A \subset [\![n_0, \infty[\![$. Se supone que A satisface las siguientes dos condiciones:*

1. $n_0 \in A$.
2. $n \in A \Rightarrow n+1 \in A$.

Entonces $A = [\![n_0, \infty[\![$.

En efecto, sea $B = \{m \in \mathbb{N} \mid m - 1 + n_0 \in A\}$. Puesto que $1 - 1 + n_0 = n_0 \in A$, $1 \in B$. Supongamos que $n \in B$, entonces $n - 1 + n_0 \in A$, y por la segunda condición que satisface este conjunto, se tiene $n + n_0 = n + 1 - 1 + n_0 \in A$; pero $n + 1 - 1 + n_0 = n + n_0 \in A$; por tanto $n + 1 \in B$. Por el principio de inducción $B = \mathbb{N}$. Ahora, si $m \in [\![n_0, \infty[\![$, entonces $m - n_0 + 1 \in \mathbb{N}$ y, por tanto, $m = m - n_0 + 1 - 1 + n_0 \in A$; esto significa $A = [\![n_0, \infty[\![$.

Definiciones por recurrencia. El principio de inducción es muy útil para construir definiciones por recurrencia. Por ejemplo, de matemáticas básicas sabemos que si n es un entero no negativo se define el factoral de n como $n! = 1 \cdot 2 \cdot \cdots \cdot n$; sin embargo, es posible definir por recurrencia el factorial de la siguiente manera: $0! = 1$, y si n es un número entero no negativo para el cual se supone ya definido $n!$, entonces se define $(n+1)! = n!(n+1)$. Esta forma de definir es mucho más útil de lo que parece y fue utilizada en este libro para definir conceptos mucho más complejos que el factorial de un número; por ejemplo, el determinante de una matriz (cfr. apartado 2.2.1).

▶ **Ejemplo A.15** Mostrar por inducción la célebre fórmula del binomio de Newton: Si $a, b \in \mathbb{R}$ y $n \in \mathbb{N}$, entonces

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k \tag{A.4}$$

donde $\binom{n}{k}$ se llama coeficiente binomial y se define por $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$. ◀

DEMOSTRACIÓN ■ Antes de probar la fórmula (A.4), necesitamos mostrar que,

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1} \tag{A.5}$$

En efecto:

$$\begin{aligned}\binom{n}{k}+\binom{n}{k+1} &= \frac{n!}{k!(n-k)!}+\frac{n!}{(k+1)!(n-k-1)!}\\ &= \frac{n!(k+1)+n!(n-k)}{(k+1)!(n-k)!}\\ &= \frac{n!(n+1)}{(k+1)!(n-k)!}\\ &= \binom{n+1}{k+1}.\end{aligned}$$

Si $n=1$,

$$\begin{aligned}(a+b)^1 &= a+b \quad \text{y}\\ \sum_{k=0}^{1}\binom{1}{k}a^{1-k}b^k &= \binom{1}{0}a+\binom{1}{1}b\\ &= a+b;\end{aligned}$$

por tanto la relación (A.4) es válida para $n=1$. Sea $n\in\mathbb{N}$ y supongamos que (A.4) es cierta para este número natural. Entonces

$$\begin{aligned}(a+b)^{n+1} &= (a+b)^n(a+b)\\ &= \left(\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k\right)(a+b);\end{aligned}$$

luego

$$(a+b)^{n+1}=\sum_{k=0}^{n}\binom{n}{k}a^{n-k+1}b^k+\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1} \tag{A.6}$$

Hagamos $j=k+1$ en la sumatoria del segundo término en el lado derecho de la precedente igualdad, entonces $k=j-1$ y, como $0\le k\le n$, se tiene $1\le j\le n+1$; por tanto

$$\begin{aligned}\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1} &= \sum_{j=1}^{n+1}\binom{n}{j-1}a^{n-j+1}b^j\\ &= \sum_{k=1}^{n+1}\binom{n}{k-1}a^{n-k+1}b^k.\end{aligned}$$

Al sustituir este último resultado en (A.6) obtenemos

$$\begin{aligned}(a+b)^{n+1} &= \sum_{k=0}^{n}\binom{n}{k}a^{n-k+1}b^k+\sum_{k=1}^{n+1}\binom{n}{k-1}a^{n-k+1}b^k\\ &= \binom{n}{0}a^{n+1}+\sum_{k=1}^{n}\binom{n}{k}a^{n-k+1}b^k+\sum_{k=1}^{n}\binom{n}{k-1}a^{n-k+1}b^k+\binom{n+1}{n}b^{n+1}\\ &= a^{n+1}+\sum_{k=1}^{n}\left(\binom{n}{k-1}+\binom{n}{k}\right)a^{n+1-k}b^k+b^{n+1}\end{aligned}$$

Luego, por (A.5),

$$\begin{aligned}(a+b)^{n+1} &= a^{n+1}+\sum_{k=1}^{n}\left(\binom{n}{k-1}+\binom{n}{k}\right)a^{n+1-k}b^k+b^{n+1}\\ &= a^{n+1}+\sum_{k=1}^{n}\binom{n+1}{k}a^{n+1-k}b^k+b^{n+1}\\ &= \sum_{k=0}^{n}\binom{n+1}{k}a^{n+1-k}b^k.\end{aligned}$$

Así, por el principio de inducción, (A.4) es válida para todo $n\in\mathbb{N}$. ■

Con frecuencia, al usar inducción, una vez que ya se probó que es verdadera para $n=1$, es necesario suponer que una proposición no únicamente es cierta para n, sino para todos los enteros $1,2\ldots,n$, y entonces probarla para el entero $n+1$. Éste es el llamado principio de inducción completa que enunciamos y probamos, como consecuencia del principio de inducción, a continuación:

Teorema A.11 **(Principio de inducción completa)** *Sea $A\subset\mathbb{N}$ un conjunto que satisface las siguientes dos condiciones:*

(i) $1\in A$.
(ii) $1,2,3,\ldots,n\in A$ *implica* $n+1\in A$.

Entonces $A=\mathbb{N}$.

DEMOSTRACIÓN ■ Sea $B=\{n\in\mathbb{N}\mid 1,2,3,\ldots,n\in A\}$. Entonces $1\in A$ por la primera condición que cumple este conjunto. Si $n\in B$, $1,2,\ldots,n\in A$, luego, por la segunda condición, $n+1\in A$ y, por tanto, $1,2,\ldots,n,n+1\in A$; entonces $n+1\in B$. Por el principio de inducción $B=\mathbb{N}$ y, en consecuencia, $A=\mathbb{N}$. ■

Obviamente el principio de inducción completa tiene su versión con corrimiento como el principio de inducción simple; la demostración es completamente análoga al caso de inducción simple con corrimiento y se deja de ejercicio al lector.

Principio de inducción completa (con corrimiento): *Sean $n_0\in\mathbb{Z}$ y $A\subset[\![n_0,\infty[\![$. Si el conjunto A satisface las siguientes condiciones:*

1. $n_0\in A$,
2. $[\![n_0,n]\!]\subset A\Rightarrow n+1\in A$,

entonces $A=[\![n_0,\infty[\![$.

► **Ejemplo A.16** **(Teorema fundamental de la aritmética)** Ahora vamos a probar, utilizando inducción completa, el llamado teorema fundamental de la aritmética, que afirma que todo entero mayor o igual a 2 se puede factorizar como el producto de números primos; esto es: *Si $m\geq 2$ es cualquier número entero, entonces existen $p_1,p_2,\ldots,p_r$ números primos tales que $m=p_1p_2\cdots p_r$.* En efecto, si $m=2$, entonces $m=p_1$, con $p_1=2$. Sea $m>2$ un entero y supongamos que la afirmación es cierta para todos los enteros $2,3,\ldots,m-1$. Si m es primo, entonces $m=p_1$ con $p_1=m$; si m no es primo, tiene divisores distintos de él y la unidad, por tanto, existen dos enteros a,b, con $2\leq a\leq b\leq m-1$, tales que $m=ab$.

Por hipótesis de inducción existen $p_1, p_2, \ldots, p_r, q_1, q_2, \ldots, q_s$ números primos tales que $a = p_1 p_2 \cdots p_r$ y $b = q_1 q_2 \cdots q_s$; luego $m = ab = p_1 p_2 \cdots p_r \; q_1 q_2 \cdots q_s$. Por el principio de inducción completa (con corrimiento) la afirmación es cierta para todo entero $m \geq 2$.◄

❍ **Nota A.3** Por comodidad redactamos los principios de inducción simple, completa con y sin corrimiento, en notación de conjuntos; pero, como es natural, se pueden traducir fácilmente al lenguaje de afirmaciones $p(n)$ como lo hicimos al inicio de esta sección. De hecho, observe que en las demostraciones de los ejemplos que realizamos utilizamos ambos lenguajes implícitamente; y es así como se procede en la práctica.

Al comienzo de este apartado utilizamos la conjetura de Goldbach (y un ejemplo previo) para ilustrar el porqué la inducción empírica no se puede utilizar para mostrar proposiciones en matemáticas de acuerdo con el método deductivo en el cual se basa esta ciencia. Sin embargo, cabe aclarar al lector que no lo sepa, la conjetura de Goldbach no se ha podido demostrar matemáticamente desde que C. Goldbach la planteó. Curiosamente, esta conjetura ha llamado la atención del público lego; de hecho se hicieron una novela y una película motivadas en el tema; incluso se ofreció un premio de un millón de dólares a quien lograra probarla rigurosamente, como ardid publicitario para promover el libro. Evidentemente la casa editora estaba bien asesorada y sabían que este problema, a pesar de tanta evidencia empírica a favor, tenía pocas posibilidades de ser resuelto, sobre todo en el lapso tan corto en el que se ofreció la recompensa.[8]

[8] N. de E. La novela es *El tío Petros y la conjetura de Goldbach*, de Apostolos Doxiadis (Madrid, Punto de Lectura, 2001) y *La habitación de Fermat* (Dir. Luis Piedrahita y Rodrigo Sopeña, España, 2007).

B Números complejos, campos y espacios vectoriales

B.1 Números complejos

Definición B.1 *Se denota por $\mathbb{C}$ al conjunto de pares ordenados (a,b), con a y b números reales; es decir,*

$$\mathbb{C} = \{(a,b) \mid a,b \in \mathbb{R}\}.$$

En $\mathbb{C}$ se definen las siguientes dos operaciones:

1. **Suma**: $(a,b)+(c,d) = (a+c, b+d)$.
2. **Producto**: $(a,b)(c,d) = (ac-bd, ad+bc)$.

Al conjunto $\mathbb{C}$, junto con estas dos operaciones, se le llama el sistema (o campo) de los **números complejos**.

La estructura algebraica del campo $\mathbb{C}$ es de mucho mayor interés que su estructura vectorial; por esta razón es que a los elementos de $\mathbb{C}$ se les dice números complejos y se les representa con notación de escalares en lugar de denotarlos como vectores. Esto no debe extrañar al lector; pues es exactamente el mismo caso que el de los números reales, los cuales también son vectores.

► **Ejemplo B.1**

- Hallar la suma de los números complejos $z_1 = (-2,5)$, $z_2 = (-4,-2)$:

$$z_1 + z_2 = (-2,5) + (-4,-2) = (-6,3).$$

- Encontrar el producto de los números complejos $z_1 = (-4,-1)$, $z_2 = (-3,2)$:

$$\begin{aligned} z_1 z_2 &= (-4,-1)(-3,2) \\ &= ((-4)(-3)-(-1)(2), (-4)(2)+(-1)(-3)) \\ &= (12+2, -8+3) \\ &= (14,-5). \blacktriangleleft \end{aligned}$$

Observe que la suma y producto de números complejos da como resultado nuevamente un número complejo. Geométricamente $\mathbb{C}$ es el plano de pares ordenados $\mathbb{R}^2$, como se ilustra en la figura B-1; sin embargo, al eje de las abscisas se le llama ahora **eje real** y al de las ordenadas **eje imaginario** (más adelante se verá la razón de esta jerga). Así, la suma de los números complejos (a,b) y (c,d), no es

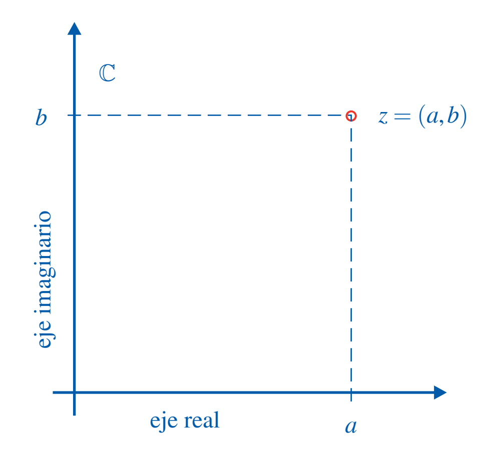


Figura B-1 •

más que la suma de los vectores $\vec{A} = (a,b)$ y $\vec{B} = (c,d)$ (cfr. figura 3-2, pág. 114). La interpretación del producto de dos números complejos requiere de un poco más de formulación, la cual daremos más adelante.

Igualdad de números complejos: Claramente, si $z_1 = (a,b) \in \mathbb{C}$ y $z_2 = (c,d) \in \mathbb{C}$, entonces $z_1 = z_2 \Leftrightarrow a = c$ y $b = d$.

Inclusión de $\mathbb{R}$ en $\mathbb{C}$: Todas las parejas de la forma $(x,0)$ las identificamos con el número real x y viceversa; es decir,

$$(x,0) =: x.$$

En particular, $(0,0) =: 0$ y $(1,0) =: 1$. De esta manera, el conjunto de los números reales queda incluido algebraicamente en el sistema de los números complejos y podemos considerar al número real x como el número complejo $(x,0)$.

El número complejo, o unidad imaginaria, i: Al número $(0,1)$ lo denotamos como i; es decir,

$$i =: (0,1)$$

y, entonces,

$$i^2 = -1\,,$$

el número complejo $(-1,0)$; pues

$$\begin{aligned} i^2 &= (0,1)(0,1) \\ &= (0\cdot 0 - 1\cdot 1, 0\cdot 1 + 1\cdot 0) \\ &= (-1,0) \\ &= -1. \end{aligned}$$

Propiedades de campo del sistema de números complejos: Es fácil verificar que el sistema de los números complejos satisface las siguientes propiedades:

Teorema B.1

1. *$\forall z_1, z_2 \in \mathbb{C}$: $z_1 + z_2 \in \mathbb{C}$ (la suma es cerrada).*
2. *$\forall z_1, z_2, z_3 \in \mathbb{C}$: $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$ (asociatividad).*
3. *$\forall z_1, z_2 \in \mathbb{C}$: $z_1 + z_2 = z_2 + z_1$ (conmutatividad).*
4. *$\forall z \in \mathbb{C}$: $z + (0,0) = z$ (existencia del neutro aditivo).*
5. *Dado $z \in \mathbb{C}$ existe $-z \in \mathbb{C}$ tal que $z + (-z) = (0,0)$ (existencia del inverso aditivo). De hecho, si $z = (a,b)$, $-z = (-a,-b)$.*
6. *$\forall z_1, z_2 \in \mathbb{C}$: $z_1 z_2 \in \mathbb{C}$ (la multiplicación es cerrada).*
7. *$\forall z_1, z_2, z_3 \in \mathbb{C}$: $z_1(z_2 z_3) = (z_1 z_2) z_3$ (asociatividad).*
8. *$\forall z_1, z_2 \in \mathbb{C}$: $z_1 z_2 = z_2 z_1$ (conmutatividad).*
9. *$\forall z \in \mathbb{C}$: $z(1,0) = z$ ((1,0) es el neutro multiplicativo de $\mathbb{C}$)*
10. *$\forall z \in \mathbb{C} - \{(0,0)\}$: existe $z^{-1} \in \mathbb{C}$ tal que $zz^{-1} = (1,0)$ (existencia del inverso multiplicativo)*
11. *$\forall z_1, z_2, z_3 \in \mathbb{C}$: $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$ (distributividad)*

Forma normal de un número complejo: Con lo anterior, si $a, b \in \mathbb{R}$, entonces

$$\begin{aligned} a + bi &= (a,0) + (b,0)(0,1) \\ &= (a,0) + (0,b) \\ &= (a,b). \end{aligned}$$

Esto es,

$$(a,b) = a + bi$$

que es la **forma normal** del número complejo $z = (a,b)$. Podemos entonces escribir

$$\mathbb{C} = \{z = a + bi \,|\, a, b \in \mathbb{R}, i^2 = -1\}$$

y operar los números complejos como si "fueran" números reales, utilizando el hecho de que $i^2 = -1$ y las propiedades de campo contenidas en el teorema B.1.

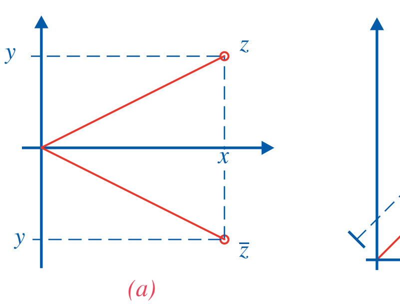
▶ **Ejemplo B.2** Si $z_1 = -3 + 2i$, $z_2 = -5 + 5i$, entonces:

1. $$\begin{aligned} z_1 + z_2 &= (-3 + 2i) + (-5 + 5i) \\ &= (-3 - 5) + (2 + 5)i \\ &= -8 + 7i. \end{aligned}$$

2. $$\begin{aligned} -z_1 &= -(-3 + 2i) \\ &= 3 - 2i. \end{aligned}$$

3. $$\begin{aligned} z_1 z_2 &= (-3 + 2i)(-5 + 5i) \\ &= 15 - 10i - 15i + 10i^2 \\ &= 15 - 25i - 10 \\ &= 5 - 25i. \blacktriangleleft \end{aligned}$$

Parte real y parte imaginaria de un número complejo: Si $z = a + ib$ se define la **parte real** de z como $\operatorname{Re}(z) = a$; y la **parte imaginaria** de z como $\operatorname{Im}(z) = b$.



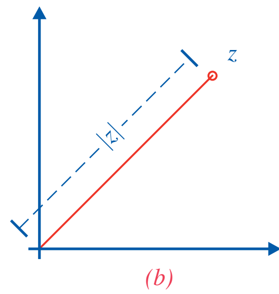


Figura B-2 •

▶ **Ejemplo B.3** Si $z = -3 + \sqrt{2}i$, se tiene: $\operatorname{Re}(z) = -3$, $\operatorname{Im}(z) = \sqrt{2}$.◀

Números imaginarios puros: Un número complejo z es **imaginario puro** si es de la forma $z = bi$, $b \neq 0$; es decir, si $\operatorname{Re}(z) = 0$ e $\operatorname{Im}(z) \neq 0$. Ya que los números imaginarios puros se encuentran en el eje de las ordenadas, a éste se le dice eje imaginario.

Conjugado de un número complejo: Si $z = a + ib$ se define y denota el **conjugado** de z como $\bar{z} = a - ib$. El conjugado de z se obtiene, geométricamente, reflejando este punto sobre el eje real como se ilustra en la figura B-2 *(a)*.

Módulo o valor absoluto de un número complejo: Si $z = a + ib$ se define el **módulo** de z, también llamado **valor absoluto** de z, como $|z| = \sqrt{a^2 + b^2}$; es decir, $|z|$ no es otra cosa que la norma del vector (a,b) (cfr. figura B-2 *(b)*). Notemos que si $z = a + ib$, entonces

$$\begin{aligned} z\bar{z} &= (a+ib)(a-ib) \\ &= a^2 - b^2 i \\ &= a^2 + b^2; \end{aligned}$$

es decir,

$$z\bar{z} = |z|^2.$$

División de números complejos: Sean $z_1, z_2 \in \mathbb{C}$, $z_2 \neq 0$, se define la división de z_1 con z_2 como

$$\frac{z_1}{z_2} = z_1 z_2^{-1}.$$

Notemos que

$$\frac{z_1}{z_2} = \frac{z_1\bar{z}_2}{z_2\bar{z}_2} = \frac{z_1\bar{z}_2}{|z_2|^2}.$$

Es decir,

$$\frac{z_1}{z_2} = \frac{z_1\bar{z}_2}{|z_2|^2}.$$

▶ **Ejemplo B.4** Calcular $\dfrac{2-3i}{1-i}$:

$$\frac{2-3i}{1-i} = \frac{(2-3i)(1+i)}{(1-i)(1+i)} = \frac{5-i}{|1-i|^2} = \frac{5-i}{2} = \frac{5}{2} - \frac{1}{2}i.◀$$

Potencias de números complejos: Sea $z \in \mathbb{C}$ y $n \in \mathbb{N}$, se definen

$$\begin{aligned} z^0 &= 1 \\ z^1 &= z \\ z^2 &= zz \\ &\vdots \\ z^n &= \underbrace{z \cdots z}_{n} \end{aligned}$$

y, si $z \neq 0$,

$$z^{-n} = \frac{1}{z^n}.$$

▶ Ejemplo B.5

$$\begin{aligned} (1+i)^3 &= (1+i)^2(1+i) \\ &= (1+2i-1)(1+i) \\ &= -2+2i. \end{aligned}$$ ◀

Forma polar de un número complejo: Al número complejo $z = x+iy$ lo podemos describir también por medio de coordenadas polares (r,θ), tomando $r = |z|$. Es decir,

$$\begin{aligned} z &= x+iy \\ &= r(\cos\theta + i\operatorname{sen}\theta) \\ &= |z|(\cos\theta + i\operatorname{sen}\theta), \end{aligned}$$

donde θ es un argumento del punto $z = (x,y)$ (cfr. figura B-3). A esta descripción la llamaremos **forma polar** del número complejo z. Recordemos que θ no está unívocamente determinado, pues cualquier otro ángulo de la forma $\theta + 2k\pi$, $k \in \mathbb{Z}$, es también un argumento de z. Denotamos por $\arg(z)$ el conjunto de todos los posibles ángulos (en radianes) que junto con $r = |z|$ forman un par de coordenadas polares de $z = (x,y)$.

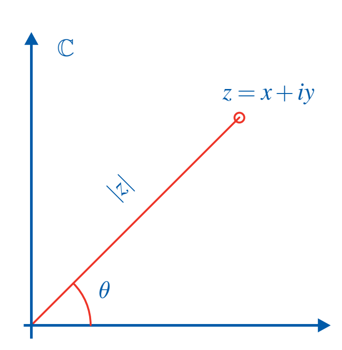


Figura B-3 •

Argumento principal de un número complejo: Si $z = a+ib$, existe un único θ, argumento de z, que satisface: $-\pi \leq \theta < \pi$. A θ se le dice **argumento principal** de z y se escribe $\operatorname{Arg}(z) = \theta$. Es fácil comprobar que si $z = x+iy$, entonces

$$\operatorname{Arg}(z) = \begin{cases} \arctan \frac{y}{x}, & \text{si } x > 0; \\ \arctan \frac{y}{x} + \pi, & \text{si } x < 0,\, y \geq 0; \\ \arctan \frac{y}{x} - \pi, & \text{si } x < 0,\, y < 0; \\ \frac{\pi}{2}, & \text{si } x = 0,\, y > 0; \\ -\frac{\pi}{2}, & \text{si } x = 0,\, y < 0. \end{cases}$$

Observe que

$$\arg z = \operatorname{Arg} z + 2k\pi, \qquad k = 0, \pm 1, \pm 2, \ldots$$

► **Ejemplo B.6** Si $z = 1 + i\sqrt{3}$, por la fórmula precedente, $\operatorname{Arg}(z) = \pi/3$ y $\arg(z) = \pi/3 + 2k\pi$, $k \in \mathbb{Z}$.◄

Interpretación geométrica del producto de complejos: Pasemos ahora a dar la interpretación geométrica del producto y el cociente de números complejos. Sean

$$z_1 = |z_1|(\cos\theta_1 + i \operatorname{sen}\theta_1),$$
$$z_2 = |z_2|(\cos\theta_2 + i \operatorname{sen}\theta_2).$$

Entonces

$$\begin{aligned} z_1 z_2 &= |z_1||z_2|(\cos\theta_1\cos\theta_2 - \operatorname{sen}\theta_1\operatorname{sen}\theta_2 + i\operatorname{sen}\theta_1\cos\theta_2 + i\cos\theta_1\operatorname{sen}\theta_2) \\ &= |z_1||z_2|(\cos\theta_1\cos\theta_2 - \operatorname{sen}\theta_1\operatorname{sen}\theta_2 + i(\operatorname{sen}\theta_1\cos\theta_2 + \cos\theta_1\operatorname{sen}\theta_2)) \end{aligned}$$

De donde,

$$z_1 z_2 = |z_1||z_2|(\cos(\theta_1 + \theta_2) + i \operatorname{sen}(\theta_1 + \theta_2)) \qquad \text{(B.1)}$$

Lo cual nos indica que el módulo del producto es el producto de los módulos: $|z_1 z_2| = |z_1||z_2|$, y que todo argumento de $z_1 z_2$ es la suma de un argumento de z_1 con un argumento de z_2; esto es,

$$\arg(z_1 z_2) = \arg z_1 + \arg z_2$$

(ver figura B-4). De manera análoga se prueba que

$$\frac{z_1}{z_2} = \frac{|z_1|}{|z_2|}(\cos(\theta_1 - \theta_2) + i \operatorname{sen}(\theta_1 - \theta_2));$$

y la interpretación geométrica es clara.

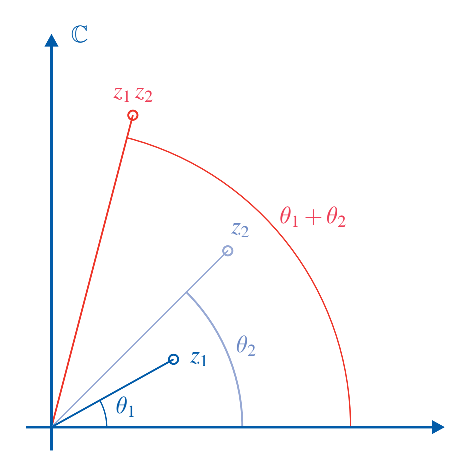


Figura B-4 •

Fórmula de De Moivre: De (B.1) se desprende inmediatamente la llamada fórmula de De Moivre para cualquier número complejo $z = r(\cos\theta + i\,\mathrm{sen}\,\theta)$ y para cualquier $n \in \mathbb{Z}$:

$$z^n = r^n(\cos(n\theta) + i\,\mathrm{sen}(n\theta)) \tag{B.2}$$

Fórmula de Euler: Para cada $\theta \in \mathbb{R}$ se define el número complejo

$$e^{i\theta} = \cos\theta + i\,\mathrm{sen}\,\theta.$$

La precedente relación es conocida como fórmula de Euler, y con ella las expresiones para el producto y la división se traducen en:

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1+\theta_2)} \quad \text{y} \quad \frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}\,.$$

Mientras que la fórmula de De Moivre, relación (B.2), se transcribe a:

$$z^n = r^n e^{in\theta} \tag{B.3}$$

El sistema de los números complejos, a diferencia del sistema de los números reales, es **algebraicamente cerrado**; es decir, todo polinomio con coeficientes en este campo tiene todas sus raíces contenidas en él. Más adelante veremos con detalle esta valiosa característica en el teorema fundamental del álgebra (cfr. teorema B.9, página 1025), esta propiedad es la razón por la que históricamente se construyeron los números complejos.

B.2 Campos

Definición B.2 *Un campo es un conjunto $\mathbb{K}$, junto con dos operaciones entre sus elementos:*

suma o adición: $a + b$,
producto o multiplicación: $a \cdot b$, *que también se representa simplemente por ab;*

que satisfacen las siguientes propiedades:

1. $a + b \in \mathbb{K}\ \ \forall a, b \in \mathbb{K}$ *(la suma es una operación cerrada).*
2. $a + (b + c) = (a + b) + c\ \ \forall a, b, c \in \mathbb{K}$ *(asociatividad de la suma).*
3. $a + b = b + a\ \ \forall a, b \in \mathbb{K}$ *(conmutatividad de la suma).*
4. *Existe* $0_{\mathbb{K}} \in \mathbb{K}$*, tal que* $a + 0_{\mathbb{K}} = a\ \ \forall a \in \mathbb{K}$ *(existencia del neutro aditivo).*
5. *Para cada* $a \in \mathbb{K}$ *existe* $-a \in \mathbb{K}$ *tal que* $a + (-a) = 0_{\mathbb{K}}$ *(existencia del inverso aditivo).*
6. $ab \in \mathbb{K}\ \ \forall a, b \in \mathbb{K}$ *(la multiplicación es una operación cerrada).*
7. $a(bc) = (ab)c\ \ \forall a, b, c \in \mathbb{K}$ *(asociatividad de la multiplicación).*
8. $ab = ba\ \ \forall a, b \in \mathbb{K}$ *(conmutatividad del producto).*
9. *Existe* $1_{\mathbb{K}} \in \mathbb{K} - \{0_{\mathbb{K}}\}$ *tal que* $a \cdot 1_{\mathbb{K}} = a\ \ \forall a \in \mathbb{K}$ *(existencia del neutro multiplicativo o identidad multiplicativa).*
10. *Para cada* $a \in \mathbb{K} - \{0_{\mathbb{K}}\}$ *existe* $a^{-1} \in \mathbb{K}$ *tal que* $aa^{-1} = 1_{\mathbb{K}}$ *(existencia del inverso multiplicativo).*
11. $a(b + c) = ab + ac\ \ \forall a, b, c \in \mathbb{K}$ *(distributividad del producto respecto a la suma).*

▶ **Ejemplo B.7**

1. $\mathbb{Q}$, el conjunto de números racionales, con las operaciones usuales, suma y multiplicación, es un campo.
2. $\mathbb{R}$, el conjunto de los números reales, con las operaciones usuales, suma y multiplicación, es un campo.
3. $\mathbb{C}$, el conjunto de los números complejos con las operaciones suma y multiplicación de la definición B.1, es un campo (cfr. propiedades del campo complejo en el teorema B.1).◀

▶ **Ejemplo B.8** **(El campo $\mathbb{Z}_2$)** Sea $\mathbb{Z}_2 = \{0,1\}$ con las operaciones suma y multiplicación definidas en las siguientes tablas:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Es fácil verificar que $\mathbb{Z}_2$ junto con estas operaciones es un campo.◀

Denotaremos por $\mathfrak{M}_{m\times p}(\mathbb{Z}_2)$ al conjunto de matrices de tamaño $m\times p$ con componentes en $\mathbb{Z}_2$; así, por ejemplo, $A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ pertenecen a $\mathfrak{M}_{2\times 3}(\mathbb{Z}_2)$; mientras que $C = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \in \mathfrak{M}_{3\times 3}(\mathbb{Z}_2)$ y, por tanto,

$$A+B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{y} \quad AC = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

En general, si n es un entero positivo, se define en el conjunto de los números enteros $\mathbb{Z}$ la relación $a \sim b$ si y sólo si n divide exactamente a $b-a$; es decir, el residuo de dividir $b-a$ entre n es cero. Se puede probar que $\sim$ es una relación de equivalencia; esto es:

1. $a \sim a \; \forall a \in \mathbb{Z}$.
2. $a \sim b \Leftrightarrow b \sim a \; \forall a,b \in \mathbb{Z}$.
3. $a,b,c \in \mathbb{Z}$, $a \sim b$, $b \sim c \Rightarrow a \sim c$.

La relación de equivalencia $\sim$ se llama relación de congruencia módulo n y si $a \sim b$ se escribe $a \equiv b \bmod n$ (se lee: "a es congruente con b módulo n"). Si representamos por $[a]$ al conjunto de todos los elementos $b \in \mathbb{Z}$ que son equivalentes al entero a; es decir, $[a] = \{b \in \mathbb{Z} \,|\, a \equiv b \bmod n\}$, se dice que este conjunto es una *clase de equivalencia*. Entonces, dadas dos clases de equivalencia, $[a]$ y $[b]$, son disjuntas o coinciden y la reunión de todas ellas resulta ser el conjunto $\mathbb{Z}$. El conjunto de todas las clases de equivalencia por la relación de congruencia módulo n se representa por el símbolo $\mathbb{Z}_n$. Si $[a] \in \mathbb{Z}_n$, por el algoritmo de división para los números enteros, existen únicos enteros r y q, con $0 \le r \le n-1$, tales que $a = nq + r$; de donde se desprende que $[a] = [r]$ y, por tanto,

$$\mathbb{Z}_n = \{[0],[1],[2],\ldots,[n-1]\}.$$

En $\mathbb{Z}_n$ se definen las operaciones

$$\begin{aligned} [a]+[b] &= [a+b] \\ [a][c] &= [ac]. \end{aligned}$$

Es fácil demostrar que las dos operaciones están bien definidas; es decir, los resultados son independientes de los representantes de cada clase de equivalencia. Además, se puede probar que con estas operaciones $\mathbb{Z}_n$ es un campo si y sólo si n es un número primo. Así, $\mathbb{Z}_3$ y $\mathbb{Z}_{19}$ son campos, pero $\mathbb{Z}_4$ y $\mathbb{Z}_{21}$ no lo son.

► **Ejemplo B.9** En el campo $\mathbb{Z}_5$, por ejemplo, $21 = 4 \cdot 5 + 1$, por tanto $[21] = [1]$. Las tablas para las operaciones suma y multiplicación son:

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

donde, para simplificar notaciones, hemos escrito simplemente k en lugar de $[k]$ para la clase de equivalencia de los enteros que son congruentes con k módulo 5. Entonces, por ejemplo, el inverso aditivo en este campo de 3 es 2 ($-[3] = [2]$) y es también su inverso multiplicativo ($[3]^{-1} = [2]$).◄

Teorema B.2 ***(Propiedades básicas de los campos)*** *Sea $\mathbb{K}$ un campo, entonces:*

1. *Los elementos (i) neutro aditivo, (ii) neutro multiplicativo, (iii) inverso aditivo e (iv) inverso multiplicativo, son únicos.*
2. *Si $a, b, c \in \mathbb{K}$:*
   i. *$a + b = a + c \Rightarrow b = c$ (ley de cancelación para la suma).*
   ii. *$ab = ac$, $a \neq 0 \Rightarrow b = c$ (ley de cancelación para el producto).*
3. $a \cdot 0_{\mathbb{K}} = 0_{\mathbb{K}} \quad \forall a \in \mathbb{K}$.
4. $-(-a) = a \quad \forall a \in \mathbb{K}$.
5. $(-a)b = a(-b) = -(ab) \quad \forall a, b \in \mathbb{K}$.
6. $(-a)(-b) = ab \quad \forall a, b \in \mathbb{K}$.
7. $(-1_{\mathbb{K}})a = -a \quad \forall a \in \mathbb{K}$.
8. *Si $a \in \mathbb{K} - \{0_{\mathbb{K}}\}$, $a^{-1} \neq 0_{\mathbb{K}}$ y $(a^{-1})^{-1} = a$.*
9. *Si $a, b \in \mathbb{K}$, $ab = 0_{\mathbb{K}} \Rightarrow a = 0_{\mathbb{K}}$ o $b = 0_{\mathbb{K}}$.*

DEMOSTRACIÓN ■ 1. *(i)* Sea $\phi \in \mathbb{K}$ tal que $a + \phi = a$ para todo $a \in \mathbb{K}$; entonces, en particular,

$$\phi = 0_{\mathbb{K}} + \phi = 0_{\mathbb{K}}\,.$$

(ii) Sean $u \in \mathbb{K}$ tal que $au = a$ para todo $a \in \mathbb{K}$, entonces, en particular,

$$1_{\mathbb{K}} = 1_{\mathbb{K}} u = u.$$

(iii) Sean $a \in \mathbb{K}$ y $g \in \mathbb{K}$ tal que $a + g = 0_{\mathbb{K}}$, entonces

$$\begin{aligned} -a + a = 0_{\mathbb{K}} &\Rightarrow (-a + a) + g = 0_{\mathbb{K}} + g \\ &\Rightarrow -a + (a + g) = g \\ &\Rightarrow -a + 0_{\mathbb{K}} = g \\ &\Rightarrow -a = g. \end{aligned}$$

(iv) Sean $a \in \mathbb{K} - \{0_{\mathbb{K}}\}$ y $h \in \mathbb{K}$ tal que $ah = 1_{\mathbb{K}}$, entonces

$$\begin{aligned} ah = 1_{\mathbb{K}} &\Rightarrow a^{-1}(ah) = a^{-1}1_{\mathbb{K}} \\ &\Rightarrow (a^{-1}a)h = a^{-1} \\ &\Rightarrow 1_{\mathbb{K}}h = a^{-1} \\ &\Rightarrow h = a^{-1}. \end{aligned}$$

2. *(i)*

$$\begin{aligned} a+b = a+c &\Rightarrow -a+(a+b) = -a+(a+c) \\ &\Rightarrow (-a+a)+b = (-a+a)+c \\ &\Rightarrow 0_{\mathbb{K}}+b = 0_{\mathbb{K}}+c \\ &\Rightarrow b = c. \end{aligned}$$

(ii) Como $a \neq 0_{\mathbb{K}}$, existe a^{-1} tal que $a^{-1}a = 1_K$, entonces

$$\begin{aligned} ab = ac &\Rightarrow a^{-1}(ab) = a^{-1}(ac) \\ &\Rightarrow (a^{-1}a)b = a^{-1}(ac) \\ &\Rightarrow 1_{\mathbb{K}}b = 1_{\mathbb{K}}c \\ &\Rightarrow b = c. \end{aligned}$$

3. Sea $a \in \mathbb{K}$, entonces, por la ley de cancelación para la suma,

$$\begin{aligned} a \cdot 0_{\mathbb{K}} + 0_{\mathbb{K}} &= a \cdot 0_{\mathbb{K}} \\ &= a(0_{\mathbb{K}} + 0_{\mathbb{K}}) \\ &= a \cdot 0_{\mathbb{K}} + a \cdot 0_{\mathbb{K}} \end{aligned}$$

implica

$$0_{\mathbb{K}} = a \cdot 0_{\mathbb{K}}.$$

4. Si $a \in \mathbb{K}$, entonces

$$(-a)+a = a+(-a) = 0_{\mathbb{K}},$$

y, por la unicidad del inverso aditivo, se desprende $-(-a) = a$.

5. Sean $a, b \in \mathbb{K}$, entonces, por la unicidad del inverso aditivo,

$$\begin{aligned} (-a)b + ab &= (-a+a)b \\ &= 0_{\mathbb{K}} \cdot b \\ &= 0_{\mathbb{K}} \end{aligned}$$

implica

$$-(ab) = (-a)b.$$

De manera análoga,

$$\begin{aligned} ab + a(-b) &= a(b+(-b)) \\ &= a \cdot 0_{\mathbb{K}} \\ &= 0_{\mathbb{K}} \end{aligned}$$

implica

$$-(ab) = a(-b).$$

6. Sean $a, b \in \mathbb{K}$, entonces, por el inciso anterior y el inciso 4,

$$\begin{aligned}(-a)(-b) &= -(a(-b))\\ &= -(-(ab))\\ &= ab.\end{aligned}$$

7. Sea $a \in \mathbb{K}$, entonces, por el inciso 5,

$$\begin{aligned}(-1_{\mathbb{K}})a &= -(1_{\mathbb{K}}a)\\ &= -(a)\\ &= -a.\end{aligned}$$

8. Sea $a \in \mathbb{K} - \{0_{\mathbb{K}}\}$, entonces

$$a^{-1}a = 1_{\mathbb{K}}$$

implica $a^{-1} \neq 0_{\mathbb{K}}$ (pues en caso contrario, por el inciso 3, se tendría $0_{\mathbb{K}} = 1_{\mathbb{K}}$); y, por la unicidad del inverso multiplicativo, se tiene $(a^{-1})^{-1} = a$.

9. Si $a \neq 0_{\mathbb{K}}$ y $ab = 0_{\mathbb{K}}$, entonces, por el inciso 3,

$$\begin{aligned}ab &= 0_{\mathbb{K}}\\ &= a \cdot 0_{\mathbb{K}}\end{aligned}$$

y, por la ley de cancelación para el producto, lo precedente implica

$$b = 0_{\mathbb{K}}. \quad \blacksquare$$

B.3 Polinomios sobre campos

B.3.1 Propiedades

Éste es un sumario de los principales conceptos de polinomios sobre campos. La mayoría de los resultados son generalizaciones directas de las propiedades de los polinomios con coeficientes reales que seguramente el lector conoce.

En lo que sigue $\mathbb{K}$ es un campo; al neutro aditivo y a la identidad multiplicativa las denotaremos simplemente por 0 y 1, respectivamente.

Definición B.3 *Sea $\mathbb{K}$ un campo. Un polinomio con coeficientes en $\mathbb{K}$ es una expresión de la forma* [1]

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

donde $n \in \mathbb{N}$, $a_j \in \mathbb{K}$ para todo $j = 0, 1, \ldots, n$.

1. *Los a_k son los coeficientes del polinomio $p(x)$.*
2. *Si $a_n \neq 0$, se dice que $p(x)$ es de* ***grado*** *n y se escribe $\operatorname{grad}(p(x)) = n$. En tal caso a a_n se le llama el* ***coeficiente principal*** *de $p(x)$; en particular, si $a_n = 1$, se dice que el polinomio es* ***mónico****.*
3. *Si en $p(x)$ los coeficientes a_k, $k = 1, \ldots, n$, son nulos, se dice que $p(x)$ es un polinomio constante.*
4. *El grado del polinomio constante $\theta(x) = 0$ se considera indefinido.*
5. *Al conjunto de todos los polinomios con coeficientes en $\mathbb{K}$, se le denota por el símbolo $\mathbb{K}[x]$. A la literal x se le dice* indeterminada *o* variable.

❍ Nota B.1

1. Por supuesto, la literal con la que se denota la indeterminada puede ser sustituida por cualquier otra literal, por ejemplo t, y los conjunto $\mathbb{K}[x]$ y $\mathbb{K}[t]$ representan entonces exactamente lo mismo.
2. Si $p(x) \in \mathbb{K}[x]$, escribiremos, con frecuencia, simplemente p para denotar al polinomio $p(x)$.
3. Se define $x^0 = 1$, entonces el polinomio $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$, se puede escribir como $p(x) = \sum_{k=0}^{n} a_k x^k$.

Definición B.4

1. *Dos polinomios $p(x), q(x) \in \mathbb{K}[x]$ de grado n, $p(x) = \sum_{k=0}^{n} a_k x^k$, $q(x) = \sum_{k=0}^{n} b_k x^k$, son iguales, y se escribe $p(x) = q(x)$, si y sólo si $a_k = b_k$ para todo $k = 0, 1, \ldots, n$.*
2. *Sea $p(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{K}[x]$, entonces $p = \theta$, el polinomio constante cero, si y sólo si $a_k = 0$ para todo $k = 0, 1, 2, \ldots, n$.*

Definición B.5 *Si $p(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{K}[x]$ y $\alpha \in \mathbb{K}$, a*

$$p(\alpha) = a_0 + a_1 \alpha + a_2 \alpha^2 + \cdots + a_n \alpha^n \in \mathbb{K}$$

se le llama la **evaluación** *del polinomio p en α.*

[1] Una manera más rigurosa de establecer este concepto es: un polinomio p es una sucesión (a_ν) de números $a_\nu \in \mathbb{K}$ cuyos términos son todos iguales a cero salvo un número finito de ellos. El mayor índice para el cual se tiene $a_n \neq 0$ y $a_{n+k} = 0$, $k = 1, 2, \ldots$, es el grado del polinomio.

Definición B.6 *Sea* $p(x) = \sum_{k=0}^{n} a_k x^k$ *y* $q(x) \sum_{k=0}^{m} a_k x^k$ *un par de polinomios en* $\mathbb{K}[x]$. *Se definen las operaciones:*

Suma: $p(x) + q(x) = \sum_{k=0}^{\text{máx}(m,n)} (a_k + b_k)x^k$; *donde* $a_{n+j} = 0$, $j = 1, \ldots, m-n$, *si* $n < m$ *y* $b_{m+j} = 0$, $j = 1, \ldots, n-m$, *si* $m < n$.

Producto: $p(x)q(x) = \sum_{k=0}^{m+n} c_k x^k$ *donde, para cada* $k = 0, 1, \ldots, m+n$, $c_k = \sum_{j=0}^{k} a_j b_{k-j}$; *es decir,* $c_k = \sum_{j+i=k} a_j b_i$.

Observe que si $p, q \in \mathbb{K}[x]$ y $\text{grad}(f) = n$, $\text{grad}(g) = m$, entonces $\text{grad}(fg) = m+n$; y si uno de ellos es el polinomio constante cero, entonces el producto es el polinomio constante cero que tiene grado indefinido.

$\mathbb{K}[x]$ es un **dominio entero**; es decir:

1. $p, q \in \mathbb{K}[x] \Rightarrow p+q, pq \in \mathbb{K}[x]$.
2. $p, q, r \in \mathbb{K}[x] \Rightarrow p + (q+r) = (p+q) + r$, $p(qr) = (pq)r$
3. $p, q \in \mathbb{K}[x] \Rightarrow p+q = q+p$, $pq = qp$.
4. Existe un polinomio $\theta \in \mathbb{K}[x]$, a saber θ es el polinomio constante cero, tal que $p + \theta = p \; \forall p \in \mathbb{K}[x]$.
5. Dado $p \in \mathbb{K}[x]$, existe $-p \in \mathbb{K}[x]$ tal que $p + (-p) = \theta$.
6. Para todo $p \in \mathbb{K}[x]$, $1p(x) = p(x)$ donde 1 representa el polinomio constante 1.
7. $p, q, r \in \mathbb{K}[x] \Rightarrow p(q+r) = pq + pr$.
8. Si $p(x), q(x) \in \mathbb{K}[x]$ y $p(x)q(x) = \theta$, θ el polinomio constante cero, entonces $p = \theta$ o $q = \theta$.

Teorema B.3 ***(Algoritmo de división)*** *Sean* $f(x), g(x) \in \mathbb{K}[x]$, *con* $g(x)$ *distinto del polinomio constante cero. Entonces existen únicos* $q(x), r(x) \in \mathbb{K}[x]$ *tales que*

$$f(x) = g(x)q(x) + r(x)$$

con $r(x) = \theta(x) = 0$, *el polinomio constante cero, o* $\text{grad}(r) < \text{grad}(g)$.

A los polinomios q y r del teorema anterior se les llama, respectivamente, el **cociente** y el **residuo**, de dividir el polinomio f entre el polinomio g.

Definición B.7 *Sean* $f, g \in \mathbb{K}[x]$, *con* $g \neq \theta$. *Se dice que el polinomio* g *divide al polinomio* f, *y se escribe* $g \mid f$, *si existe* $q \in \mathbb{K}[x]$ *tal que* $f = gq$. *Esto es, si* $r = \theta$ *en el algoritmo de división.*

Definición B.8 *Sean* $f, g \in \mathbb{K}[x]$; un máximo común divisor *de* f *y* g *es un polinomio* $d \in \mathbb{K}[x] - \{\theta\}$, *tal que* $d \mid f$ *y* $d \mid g$; *y si* d' *es otro polinomio en* $\mathbb{K}[x]$ *que divide a* f *y a* g, *entonces* $d' \mid d$.

Teorema B.4 *Sea $f, g \in \mathbb{K}[x]$ un par de polinomios no ambos nulos. Entonces existe al menos un máximo común divisor de f y g.*

Es evidente que si d es un máximo común divisor de f y g y $c \in \mathbb{K} - \{0\}$, entonces cd también es un máximo común divisor de f y g. Luego, si f y g son polinomios, no ambos nulos, existe un único polinomio mónico que es máximo común divisor de f y g; a éste se le llama **el máximo común divisor** de f y g y se le denota por

$$\operatorname{mcd}(f,g)$$

o por $(f(x), g(x))$.

Definición B.9 *Si* $\operatorname{mcd}(f,g) = 1$*, se dice que f y g son **primos relativos.***

Teorema B.5 *Si $f, g \in \mathbb{K}[x]$ son dos polinomios no ambos nulos, entonces existen $q, p \in \mathbb{K}[x]$ tales que*

$$\operatorname{mcd}(f,g) = qf + pg.$$

Definición B.10 *Sean $f, g \in \mathbb{K}[x]$ y p un polinomio con coeficientes en $\mathbb{K}$; p es un mínimo común múltiplo de $f(x)$ y $g(x)$ si $f \mid p$, $g \mid p$ y si q es cualquier otro polinomio que es divisible por f y g, entonces $p \mid q$.*

Teorema B.6 *Si f y g son polinomios en $\mathbb{K}[x]$ entonces ellos tienen al menos un mínimo común múltiplo.*

Si p es un mínimo común múltiplo de f y g, también lo es cp, para cualquier $c \in \mathbb{K}$. Por tanto existe un único polinomio mónico que es mínimo común múltiplo de f y g. A éste se le representa por el símbolo

$$\operatorname{mcm}(f,g)$$

o por el símbolo $[f,g]$ y se le llama **el mínimo común múltiplo** de f y g.

Teorema B.7

$$[f,g] = \frac{fg}{(f,g)}$$

Definición B.11 *Un polinomio $f(x) \in \mathbb{K}[x]$ es **irreducible** en $\mathbb{K}[x]$ si cumple la condición:*

$$p, q \in \mathbb{K}[x],\ f = pq \Rightarrow p \text{ o } q \text{ es un polinomio constante.}$$

Teorema B.8 ***(Teorema de factorizacion única)*** *Todo polinomio no nulo $f(x) \in \mathbb{K}[x]$ se puede factorizar como producto de factores mónicos irreducibles de manera única. Es decir, existen $p_1, p_2, \ldots, p_r \in \mathbb{K}[x]$ y $c \in \mathbb{K}$, tales que*

$$f(x) = cp_1(x)p_2(x)\cdots p_r(x)$$

donde los polinomios p_j son mónicos e irreducibles. Además, salvo el orden de los factores, la factorización es única.

B.3.2 Raíces y teorema fundamental del álgebra

Definición B.12 *Sean $f(x) \in \mathbb{K}[x]$ y $\alpha \in \mathbb{K}$. Se dice que α es **raíz** del polinomio $f(x)$ si $f(\alpha) = 0$.*

Evidentemente las siguientes condiciones son equivalentes a pares:

1. α es raíz del polinomio $f(x)$.
2. $(x-\alpha) \mid f(x)$.
3. El residuo de dividir el polinomio $f(x)$ entre el polinomio $(x-\alpha)$ es el polinomio constante cero.

Como consecuencia del inciso 2, un polinomio de grado n con coeficientes en un campo $\mathbb{K}$ tiene a lo más n raíces en ese campo.

Se dice que una raíz α de un polinomio $f(x)$ tiene **multiplicidad** $m \in \mathbb{N}$, si $(x-\alpha)^m \mid f(x)$ pero $(x-\alpha)^{m+1} \nmid f(x)$; en particular, si $m = 1$, se dice que α es una **raíz simple**.

A continuación enunciamos uno de los más conspicuos teoremas que se tienen en álgebra. De hecho, la construcción del campo $\mathbb{C}$ fue diseñada con el propósito de obtener este resultado para polinomios con coeficientes reales. El lector interesado en estudiar la demostración de esta otra piedra angular de las matemáticas, puede consultar el apéndice C (teorema C.8 página 1052).

Teorema B.9 ***(Teorema fundamental del álgebra)*** *Sea $p(x)$ un polinomio con coeficientes reales y/o complejos (i.e., $p(x) \in \mathbb{C}[x]$) que tiene grado mayor o igual a uno, entonces $p(x)$ tiene por lo menos una raíz en $\mathbb{C}$.*

Corolario B.1 *Sea $p(x)$ un polinomio con coeficientes reales y/o complejos (i.e., $p(x) \in \mathbb{C}[x]$) que tiene grado $n \geq 1$, entonces $p(x)$ tiene exactamente n raíces en $\mathbb{C}$ contando multiplicidades. Es decir, existen $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{C}$, números complejos distintos entre sí, tales que*

$$p(x) = a_n(x-\alpha_1)^{m_1}(x-\alpha_2)^{m_2}\cdots(x-\alpha_r)^{m_r}$$

donde m_j es la multiplicidad de la raíz α_j, $j = 1, 2, \ldots, r$; $m_1 + m_2 + \cdots + m_r = n$ y a_n es el coeficiente principal de $p(x)$.

Un campo $\mathbb{K}$ es **algebraicamente cerrado** si todo polinomio con coeficientes en $\mathbb{K}$ tiene todas sus raíces en $\mathbb{K}$. Así, el teorema fundamental del álgebra (y su corolario) garantizan que el sistema de los números complejos $\mathbb{C}$ es algebraicamente cerrado.

El teorema fundamental del álgebra se puede extender a campos arbitrarios.

Definición B.13 *Sean $\mathbb{F}$ un campo y $\mathbb{S} \subset \mathbb{F}$. Se dice que $\mathbb{S}$ es un subcampo de $\mathbb{F}$ si las mismas operaciones, suma y producto, de $\mathbb{F}$ restringidas a $\mathbb{S}$, hacen de éste un campo.*

Definición B.14 *Sea $\mathbb{K}$ un campo. Se dice que un campo $\mathbb{F}$ es una extensión del campo $\mathbb{K}$, si $\mathbb{K}$ es un subcampo de $\mathbb{F}$.*

Teorema B.10 ***(Teorema de Kronecker)*** *Sea $p(x) \in \mathbb{K}[x]$ un polinomio de grado $n \geq 1$. Entonces existe un campo $\mathbb{F}$, extensión del campo $\mathbb{K}$, y $\alpha \in \mathbb{F}$ tal que α es raíz del polinomio $p(x)$ en $\mathbb{F}$.*

Corolario B.2 *Si $p(x) \in \mathbb{K}[x]$ es un polinomio de grado $n \geq 1$, entonces existe una extensión $\mathbb{F}$ de $\mathbb{K}$ en la que $p(x)$ tiene todas sus raíces.*

Hacemos una digresión final en este segmento con un resultado, que junto con división sintética, puede ser muy útil; pues da condiciones necesarias para que un polinomio con coeficientes en el dominio de los números enteros $\mathbb{Z}$ tenga una raíz en el campo de los números racionales $\mathbb{Q}$; está contemplado en el siguiente teorema y la demostración se puede consultar en el teorema A.9 de la página 1001.

Teorema B.11 *Sea $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$ un polinomio de coeficientes enteros y de grado $n \geq 1$. Si $\dfrac{p}{q} \in \mathbb{Q}$ es una raíz racional de este polinomio; esto es $f(p/q) = 0$, entonces $p \mid a_0$ y $q \mid a_n$.*

B.4 Espacios vectoriales sobre otros campos

Definición B.15 *Sea $\mathbb{K}$ un campo.*[2] *Un espacio vectorial sobre $\mathbb{K}$ es un conjunto $\mathbf{E} \neq \emptyset$ donde se han definido un par de operaciones: suma entre sus elementos, representada como $\vec{u} + \vec{v}$; y multiplicación (producto) de escalares (elementos del campo $\mathbb{K}$) con elementos de $\mathbf{E}$, representada por $\alpha\vec{u}$; que cumplen con las siguientes diez condiciones (axiomas de espacio vectorial):*

1. $\vec{u} + \vec{v} \in \mathbf{E} \quad \forall \vec{u}, \vec{v} \in \mathbf{E}$ *(la suma es cerrada).*
2. $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w} \quad \forall \vec{u}, \vec{v}, \vec{w} \in \mathbf{E}$ *(asociatividad de la suma).*
3. $\vec{u} + \vec{v} = \vec{v} + \vec{u} \quad \forall \vec{u}, \vec{v} \in \mathbf{E}$ *(conmutatividad de la suma).*
4. *Existe* $\vec{0}_{\mathbf{E}} \in \mathbf{E}$ *tal que* $\vec{u} + \vec{0}_{\mathbf{E}} = \vec{u} \quad \forall \vec{u} \in \mathbf{E}$ *(existencia del neutro aditivo).*
5. *Para cada* $\vec{u} \in \mathbf{E}$ *existe* $-\vec{u} \in \mathbf{E}$ *tal que* $\vec{u} + (-\vec{u}) = \vec{0}_{\mathbf{E}}$ *(existencia del inverso aditivo).*
6. $\alpha\vec{u} \in \mathbf{E} \quad \forall \alpha \in \mathbb{K}, \forall \vec{u} \in \mathbf{E}$ *(la multiplicación con escalares es cerrada).*
7. $\alpha(\beta\vec{u}) = (\alpha\beta)\vec{u} \quad \forall \alpha, \beta \in \mathbb{K}, \forall \vec{u} \in \mathbf{E}$ *(asociatividad del producto con escalares).*
8. $\alpha(\vec{u} + \vec{v}) = \alpha\vec{u} + \beta\vec{v} \quad \forall \alpha \in \mathbb{K}, \forall \vec{u}, \vec{v} \in \mathbf{E}$ *(distributividad del producto respecto a la suma de vectores).*
9. $(\alpha + \beta)\vec{u} = \alpha\vec{u} + \beta\vec{u} \quad \forall \alpha, \beta \in \mathbb{K}, \forall \vec{u} \in \mathbf{E}$ *(distributividad del producto respecto a la suma de escalares).*
10. $1_{\mathbb{K}}\vec{u} = \vec{u} \quad \forall \vec{u} \in \mathbf{E}$ *(preservación de la escala).*

[2]Cfr. definición B.2 de la sección B.2, página 1017, de este apéndice.

❍ **Nota B.2** Todos los conceptos y resultados que se establecieron en los primeros tres capítulos de este libro, salvo unas cuantas obvias excepciones,[3] siguen siendo válidos si se considera como campo subyacente de escalares un campo cualquiera $\mathbb{K}$ en lugar del campo $\mathbb{R}$ (o $\mathbb{C}$). En particular: el álgebra de matrices y matrices invertibles; la teoría de determinantes; la teoría de independencia lineal, dimensión, rango y bases; la teoría general de sistemas lineales; métodos de Gauss y Gauss-Jordan; y toda la teoría de espacios vectoriales y operadores lineales en ellos.

► **Ejemplo B.10** Sean $\mathbf{E}=\mathbb{R}$, el conjunto de números reales, y $\mathbb{K}=\mathbb{R}$, el campo de los números reales. Entonces, con la suma usual de números reales y el producto usual de números reales, $\mathbf{E}$ es un espacio vectorial sobre $\mathbb{K}=\mathbb{R}$. Es fácil ver que en este caso $\dim(\mathbf{E})=1$; pues claramente $\{1\}$ es una base para este espacio.◄

► **Ejemplo B.11** Sean n un entero positivo, $\mathbf{E}=\mathbb{Q}^n=\{(x_1,x_2,\ldots,x_n)\,|\,x_i\in\mathbb{Q},\, i=1,2,\ldots,n\}$ y $\mathbb{K}=\mathbb{Q}$, con las operaciones usuales de suma de vectores en $\mathbb{R}^n$ —restringida a $\mathbb{Q}^n$— y la multiplicación por un escalar con un vector de $\mathbb{R}^n$ —restringida a escalares racionales y vectores en $\mathbb{Q}^n$—. Es sencillo verificar que $\mathbf{E}$ es un espacio vectorial sobre $\mathbb{Q}$ y que $\dim(\mathbf{E})=n$.◄

► **Ejemplo B.12** Sean n un entero positivo y $\mathbb{K}$ un campo. Se define

$$\mathbb{K}^n=\{\vec{\zeta}=(\zeta_1,\zeta_2,\ldots,\zeta_n)\mid\zeta_j\in\mathbb{K},\ j=1,2,\ldots,n\};$$

es decir, el conjunto de n-eadas ordenadas[4] con componentes en $\mathbb{K}$. Con las operaciones:

- $(\zeta_1,\zeta_2,\ldots,\zeta_n)+(\xi_1,\xi_2,\ldots,\xi_n)=(\zeta_1+\xi_1,\zeta_2+\xi_2,\ldots,\zeta_n+\xi_n)$
- $\lambda(\zeta_1,\zeta_2,\ldots,\zeta_n)=(\lambda\zeta_1,\lambda\zeta_2,\ldots,\lambda\zeta_n)$, $\lambda\in\mathbb{K}$

$\mathbb{K}^n$ es un espacio vectorial sobre $\mathbb{K}$.◄

Antes de dar el siguiente (relativamente sorprendente) ejemplo, necesitamos de algunos preliminares:

1. Sean A y B dos conjuntos. Se dice que A y B son equipotentes si existe una biyección φ de A sobre B; es decir, $\varphi: A\to B$ es una función inyectiva y suprayectiva.[5]
2. Si A y B son equipotentes se dice que ambos tienen la misma cardinalidad y se escribe

$$\operatorname{card}(A)=\operatorname{card}(B).$$

   Es decir, A y B son equipotentes si y sólo si $\operatorname{card}(A)=\operatorname{card}(B)$.
3. Un conjunto A es infinito si y sólo si A es equipotente a un subconjunto propio de A. Por ejemplo, la función $\varphi:\mathbb{N}\to\mathbb{P}$, donde $\mathbb{N}$ es el conjunto de los números naturales y $\mathbb{P}$ el conjunto de números

[3] Por ejemplo, un sistema consistente con más incógnitas que ecuaciones no puede tener una infinidad de soluciones en un campo finito.

[4] Naturalmente si $\vec{\zeta}_1=(\zeta_1,\zeta_1,\ldots,\zeta_n)\in\mathbb{K}^n$ y $\vec{\zeta}_2=(\zeta_1',\zeta_2',\ldots,\zeta_n')\in\mathbb{K}^n$, $\vec{\zeta}_1=\vec{\zeta}_2\Leftrightarrow\zeta_j=\zeta_j'$ para todo j.

[5] Cfr. definiciones 5.3 y 5.4 del apartado 5.1.2

pares positivos, definida por $\varphi(n) = 2n$, claramente es una biyección; luego $\mathbb{N}$ y $\mathbb{P}$ son equipotentes, pero $\mathbb{P} \subsetneq \mathbb{N}$; así que $\mathbb{N}$ es un conjunto infinito.

4. Se dice que el conjunto infinito $\mathbb{N}$ tiene cardinalidad $\aleph_0$ (alph cero) y se escribe
$$\operatorname{card}(\mathbb{N}) = \aleph_0.$$
5. $A \neq \emptyset$ es finito si y sólo si A es equivalente a un conjunto de la forma $\{1, 2, \ldots, n\}$ para algún entero positivo n. En tal caso se dice que A tiene cardinalidad n y se escribe $\operatorname{card}(A) = n$. Es decir, la cardinalidad de un conjunto finito es el número de elementos del conjunto.
6. Si A y B son conjuntos finitos, son equipotentes si y sólo si tienen el mismo número de elementos.
7. La cardinalidad del conjunto vacío se define como cero; es decir, $\operatorname{card}(\emptyset) = 0$.
8. Si existe una función inyectiva $\varphi : A \to B$ que no es suprayectiva, se dice que la cardinalidad de A es menor que la cardinalidad de B y se escribe $\operatorname{card}(A) < \operatorname{card}(B)$.
9. Un conjunto A es numerable si es equipotente a $\mathbb{N}$; esto es, si $\operatorname{card}(A) = \aleph_0$.
10. Se dice que un conjunto A es a lo sumo numerable si es finito o numerable.
11. Todo subconjunto de un conjunto numerable es a lo sumo numerable.
12. Que un conjunto A sea numerable significa que sus elementos se pueden escribir como una sucesión; esto es,
$$A = \{a_1, a_2, \ldots, a_n, \ldots\}.$$
13. Si $\{A_n\}$, $n = 1, 2, \ldots$, es una colección numerable de conjuntos a lo sumo numerables, entonces
$$A = \bigcup_{n=1}^{\infty} A_n$$
es un conjunto a lo sumo numerable.
14. Se puede probar que el conjunto $\mathbb{Q}$ es a lo sumo numerable.
15. La función $\varphi : \mathbb{N} \to [0, 1]$, $\varphi(n) = 1/n$, es inyectiva, pero no suprayectiva; luego $\operatorname{card}(\mathbb{N}) < \operatorname{card}([0, 1])$. A la cardinalidad del intervalo $[0, 1]$ se le llama cardinalidad del continuo y se representa por el símbolo $\mathfrak{c}$; es decir, $\operatorname{card}([0, 1]) = \mathfrak{c}$. Por tanto $\aleph_0 < \mathfrak{c}$.
16. Se puede probar que cualquier intervalo no vacío, acotado o no, tiene también la cardinalidad del continuo; en particular el conjunto de los números reales $\mathbb{R} = (-\infty, \infty)$.
17. *Hipótesis del continuo*: Si B es un conjunto tal que $\aleph_0 \leq \operatorname{card}(B) < \mathfrak{c}$, entonces $\operatorname{card}(B) = \aleph_0$. Es decir, no existe un conjunto infinito no numerable que tenga cardinalidad inferior a la cardinalidad del continuo.
18. Si A y B son conjuntos a lo sumo numerables, entonces $A \times B$ es a lo sumo numerable. En general, si $A_1, A_2, \ldots, A_n$ son conjuntos a lo sumo numerables, entonces $A_1 \times A_2 \times \cdots \times A_n$ es a lo sumo numerable.

► **Ejemplo B.13** Sean $\mathbf{E} = \mathbb{R}$, el conjunto de números reales, y $\mathbb{K} = \mathbb{Q}$, el campo de números racionales. Entonces, con la suma usual de números reales y el producto usual de números reales, claramente $\mathbf{E}$ es un espacio vectorial sobre $\mathbb{K} = \mathbb{Q}$. Mostrar que, en este caso, $\dim(\mathbf{E}) = \mathfrak{c}$.◄

DEMOSTRACIÓN ■ En efecto, sea $\mathscr{B}$ una base de $\mathbf{E}$ —la cual existe por el teorema C.5 del apéndice C, p. 1047— entonces, puesto que $\mathscr{B} \subset \mathbb{R}$, $\text{card}(\mathscr{B}) \leq \text{card}(\mathbb{R}) = \mathfrak{c}$. Supongamos que $\text{card}(\mathscr{B}) < \mathfrak{c}$, entonces, por la hipótesis del continuo, $\text{card}(\mathscr{B}) \leq \aleph_0$; es decir, $\mathscr{B}$ es a lo sumo numerable. Por tanto, $\mathscr{B}$ se puede escribir como una sucesión (posiblemente finita) de números reales; digamos $\mathscr{B} = \{\alpha_1, \alpha_2, \ldots\}$. Se define, para cada $n \in \mathbb{N}$, el subespacio $S_n = \text{gn}(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Sea $\varphi : S_n \to \mathbb{Q}^n$ la función definida, para cada $a_1\alpha_1 + a_2\alpha_2 + \cdots + a\alpha_n \in S_n$, como $\varphi(a_1\alpha_1 + a_2\alpha_2 + \cdots + a\alpha_n) = (a_1, a_2, \ldots, a_n)$. Claramente φ es una función inyectiva de S_n en el conjunto numerable $\mathbb{Q}^n$; luego $\text{card}(S_n) \leq \aleph_0$ y, por tanto, S_n es a lo sumo numerable. Puesto que la reunión de una familia numerable de conjuntos a lo sumo numerables es también un conjunto a lo sumo numerable y evidentemente $\mathbb{R} = \bigcup_{n=1}^{\infty} S_n$, la suposición $\text{card}(\mathscr{B}) < \mathfrak{c}$ lleva a la contradicción de que $\mathbb{R}$ es a lo sumo numerable. Por lo tanto, $\text{card}(\mathscr{B}) = \mathfrak{c}$. ■

► **Ejemplo B.14** **(El espacio vectorial $\mathbb{Z}_2^n$)** Sea $\mathbb{K} = \mathbb{Z}_2$ el campo descrito en el ejemplo B.8, pág. 1018, y n un entero no negativo. Si

$$\mathbb{Z}_2^n = \{\vec{u} = (x_1, x_2, \ldots, x_n) \,|\, x_j \in \mathbb{Z}_2, j = 1, 2, \ldots n\},$$

en $\mathbb{Z}_2^n$ se definen las operaciones, para cada par de vectores $\vec{u} = (x_1, x_2, \ldots, x_n)$ $\vec{v} = (y_1, x_2, \ldots, x_n)$ de $\mathbb{Z}_2^n$ y para cada $\lambda \in \mathbb{Z}_2$,

$$\begin{aligned} \vec{u} + \vec{v} &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \quad \text{y} \\ \lambda\vec{u} &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n). \end{aligned}$$

Es fácil verificar que $\mathbb{Z}_2^n$, con estas operaciones, es un espacio vectorial sobre el campo $\mathbb{Z}_2$ y que una base para este espacio (la *base canónica*) está dada por $\mathscr{B} = \{\vec{\delta}_j \mid j = 1, 2, \ldots, n\}$ donde $\vec{\delta}_j = \underbrace{(0, \ldots, 0, \overset{j}{1}, 0, \ldots, 0)}_{n}$, para cada $j = 1, 2, \ldots, n$; y, por tanto, $\dim(\mathbb{Z}_2^n) = n$. Además, es claro que $\mathbb{Z}_2^n$ tiene 2^n elementos. ◄

► **Ejemplo B.15** Si $\vec{u} = (1, 1, 0, 0, 1)$, $\vec{v} = (0, 1, 1, 0, 1)$, $\vec{u}, \vec{v} \in \mathbb{Z}_2^5$, entonces

1. $$\begin{aligned} \vec{u} + \vec{v} &= (1, 1, 0, 0, 1) + (0, 1, 1, 0, 1) \\ &= (1, 0, 1, 0, 0). \end{aligned}$$

2. $$\begin{aligned} -\vec{v} &= -(0, 1, 1, 0, 1) \\ &= (0, -1, -1, 0, -1) \\ &= (0, 1, 1, 0, 1). \end{aligned}$$

3. Si $r \in \mathbb{Z}_2$, y $\vec{w} \in \mathbb{Z}_2^5$,

$$r\vec{u} = \begin{cases} \vec{u}, & \text{si } r = 1 \\ \vec{0}, & \text{si } r = 0 \end{cases}$$ ◄

B.5 Aplicación a la teoría de detección y corrección de errores en códigos

El espacio $\mathbb{Z}_2^n$ del ejemplo B.14 tiene una importante aplicación en la teoría de detección y corrección de errores en códigos binarios que se transmiten electrónicamente. Un código binario es un conjunto de vectores en $\mathbb{Z}_2^n$. El proceso de convertir un mensaje en vectores del espacio $\mathbb{Z}_2^n$ se llama codificación (binaria), y al proceso inverso decodificación.

Una vez que un mensaje ha sido codificado, éste debe transmitirse a través de un canal (un radiotransmisor, una línea telefónica, un cable de fibra óptica, el espacio vacío, etc.); pero, en muchos casos, el canal de transmisión está sujeto a "ruido" que resulta de otras señales de interferencia; como consecuencia el mensaje original puede presentar errores al llegar al receptor. Esto significa que algunas de las entradas de los vectores código han cambiado de unos a ceros o viceversa. Existe una técnica para poder detectar si un vector código dado contiene un error, y se llama verificación de paridad. Para aplicar esta técnica supondremos que en cualquier vector código que llega al receptor, el posible error que contiene es simple; esto es, a lo más una de sus entradas ha cambiado de cero a uno o viceversa; esto significa que la probabilidad de que en un vector código existan más de dos entradas con errores es casi nula. Supongamos, por ejemplo, que se va a transmitir un mensaje de dos palabras y en sistema decimal están representadas por los número 25 y 27. En el sistema binario estos números corresponden a 11001 y 11011, respectivamente. Se dice que cada uno de estos mensajes son palabras o vectores código de longitud 5 y, en general, de longitud k si tienen, entre ceros y unos, k componentes. Para poder detectar errores en la transmisión, se añade al código original un uno al final si el número de unos es impar y un cero si es par. De esta manera, los vectores código —con verificación de paridad— para las palabras del hipotético mensaje, quedan como se ilustra en la figura B-5.

Es evidente que si sólo una de las entradas ha sido alterada en el canal, el verificador de paridad ya no será congruente con la paridad de unos en las primeras cinco componentes y se habrá detectado que el vector código contiene un error. Sin embargo, el método de verificación de paridad, aunque detecta errores, no funciona para determinar en qué componente del vector código se encuentra el error y así poder corregirlo. Para poder alcanzar este objetivo utilizaremos la siguiente técnica, que a mediados del siglo pasado fue creada por el matemático Richard W. Hamming.

Códigos de detección y corrección de errores (n, k) de Hamming

Supongamos que se van a enviar mensajes binarios con palabras de longitud k. Sean $\vec{u}_j = (\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{(n-k)j})$, $j = 1, 2, \ldots, k$, k vectores en $\mathbb{Z}_2^{n-k}$ $(n > k)$; definimos los siguientes k vectores en $\mathbb{Z}_2^n$

$$\vec{\eta}_j = (\underbrace{0, \ldots, 0, \overset{j}{1}, 0, \ldots, 0}_{k}, \alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{(n-k)j}) = \begin{bmatrix} \vec{e}_j \\ \vec{u}_j \end{bmatrix}$$

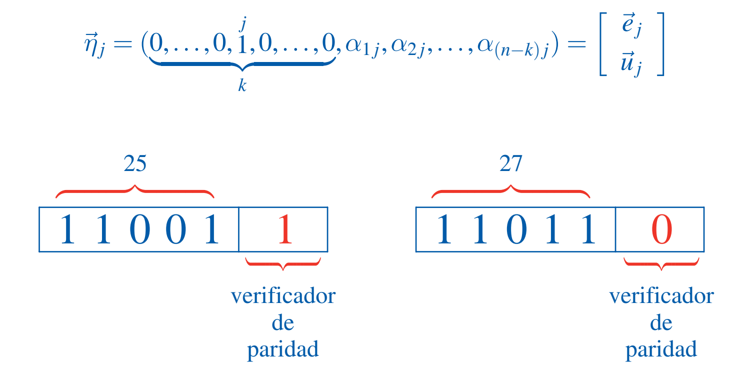


Figura B-5 •

donde $\vec{e}_j = (\underbrace{0,\ldots,0,\overset{j}{1},0,\ldots,0}_{k})$, $j = 1,2,\ldots,k$. Puesto que la matriz que tiene por filas a los vectores $\vec{\eta}_j$ está en forma escalonada y toda fila tiene pivote, los vectores $\vec{\eta}_j$ son linealmente independientes en $\mathbb{Z}_2^n$. Sean $\vec{\delta}_i = (\underbrace{0,\ldots,0,\overset{i}{1},0,\ldots,0}_{n})$ los vectores de la base canónica de $\mathbb{Z}_2^n$ y supongamos, además, que es posible construir una transformación lineal $T : \mathbb{Z}_2^n \to \mathbb{Z}_2^{n-k}$ tal que

1. $\text{Ker}(T) = \text{gn}(\vec{\eta}_1, \vec{\eta}_2, \ldots, \vec{\eta}_k)$.
2. $T(\vec{\delta}_i) \neq \vec{0}_{\mathbb{Z}_2^{n-k}}$ para todo $i = 1,2,\ldots,n$.
3. $T(\vec{\delta}_i) \neq T(\vec{\delta}_j)$ si $i \neq j$.

Sea $\vec{\omega} = (x_1, x_2, \ldots, x_k)$ una palabra binaria de longitud k que se va a transmitir por algún canal y

$$\vec{\omega}_C = \vec{x}_1\vec{\eta}_1 + x_2\vec{\eta}_2 + \cdots + x_k\vec{\eta}_k \in \text{Ker}(T),$$

entonces $\vec{\omega}_C$ es una palabra de longitud n, donde las primeras k coordenadas son idénticas a las primeras k coordenadas de $\vec{\omega}$ y las últimas coordenadas contienen la información redundante suficiente (código de verificación de paridad) para detectar y corregir errores. En efecto, si la palabra $\vec{\omega}_C'$ que recibe el receptor tiene un sólo error —digamos— en la coordenada i, entonces $\vec{\omega}_C' = \vec{\omega}_C + \vec{\delta}_i$; por tanto,

$$\begin{aligned} T(\vec{\omega}_C') &= T(\vec{\omega}_C + \vec{\delta}_i) \\ &= T(\vec{\omega}_C) + T(\vec{\delta}_i) \\ &= \vec{0}_{\mathbb{Z}_2^{n-k}} + T(\vec{\delta}_i) \\ &= T(\vec{\delta}_i) \\ &\neq \vec{0}_{\mathbb{Z}_2^{n-k}}. \end{aligned}$$

Además, si $[T]$ es la representación matricial de T relativa a las bases canónicas de $\mathbb{Z}_2^n$ y $\mathbb{Z}_2^{n-k}$, entonces $T(\vec{\delta}_i)$ es la columna i de $[T]$. De esta manera, al identificar $T(\vec{\delta}_i) = [T]\vec{\delta}_i$ con la única columna de $[T]$ con la que coincide, se puede determinar el valor i de la coordenada donde se encuentra el error y corregirlo. Por otra parte, si $\vec{\omega}_C'$ no tiene error, entonces $T(\vec{\omega}_C') = T(\vec{\omega}_C) = \vec{0}_{\mathbb{Z}_2^{n-k}}$ y viceversa. En estas condiciones se dice que el subespacio k dimensional de $\mathbb{Z}_2^n$, $C_{n,k} = \text{gn}(\vec{\eta}_1, \vec{\eta}_2, \ldots, \vec{\eta}_k)$, es un código (n,k) de Hamming para la detección y corrección de errores (simples). La matriz G que tiene por filas a los vectores $\vec{\eta}_j$ se llama matriz generadora de código de paridad y a la matriz $H = [T]$ se le dice matriz verificadora de código de paridad.

▶ **Ejemplo B.16 (Código (6,3) de Hamming)** Sean $\vec{u}_1 = (1,0,1)$, $\vec{u}_2 = (1,1,0)$, y $\vec{u}_3 = (1,1,1)$, entonces $\vec{\eta}_1 = (1,0,0,1,0,1)$, $\vec{\eta}_2 = (0,1,0,1,1,0)$ y $\vec{\eta}_3 = (0,0,1,1,1,1)$. Completemos[6] el conjunto L.I. $\{\vec{\eta}_1, \vec{\eta}_2, \vec{\eta}_3\}$ a una base del espacio $\mathbb{Z}_2^6$; el resultado que se obtiene es

$$\mathscr{B} = \{(1,0,0,0,0,0), (0,1,0,0,0,0), (0,0,1,0,0,0), \vec{\eta}_1, \vec{\eta}_2, \vec{\eta}_3\}.$$

[6] Cfr. compleción de un conjunto L.I. a una base, página 167.

Sea $T : \mathbb{Z}_2^6 \to \mathbb{Z}_2^3$ la transformación lineal tal que[7] $T(\vec{\eta}_i) = \vec{0}_{\mathbb{Z}_2^3}$, $i = 1,2,3$, $T(1,0,0,0,0,0) = (1,0,0)$ y $T(0,1,0,0,0,0) = (0,1,0)$ y $T(0,0,1,0,0,0) = (0,0,1)$. Puesto que la matriz cambió de base[8] de la base canónica a la base $\mathscr{B}$ está dada por

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

y la representación matricial de T relativa a la base $\mathscr{B}$ y a la base canónica $\{\vec{\delta}_i\}$ de $\mathbb{Z}_2^3$ es[9]

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

se tiene

$$\begin{aligned} [T] &= P \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \\ &= H. \end{aligned}$$

Por construcción, $\operatorname{Ker}(T) = \operatorname{gn}(\vec{\eta}_1, \vec{\eta}_2, \vec{\eta}_3)$; y al observar las columnas de $[T]$ se desprende que $T(\vec{\delta}_i) \neq \vec{0}_{\mathbb{Z}_2^3}$ para todo i y $T(\vec{\delta}_i) \neq T(\vec{\delta}_j)$ si $i \neq j$. En este caso las matrices generadora de código y verificadora de código de paridad están dadas por

$$G = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{y} \quad H = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix},$$

respectivamente. De esta manera, si se desea enviar, por ejemplo, la palabra de longitud 3, $\vec{\omega} = (1,1,1) = 1 \cdot \vec{\delta}_1 + 1 \cdot \vec{\delta}_2 + 1 \cdot \vec{\delta}_3$, la codificación correspondiente con código de verificación de paridad es

$$\vec{\omega}_C = 1 \cdot \vec{\eta}_1 + 1 \cdot \vec{\eta}_2 + 1 \cdot \vec{\eta}_3 = G^t \vec{\omega} = G^t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = (1,1,1,1,0,0).$$

[7] Cfr. ejercicio resuelto 1 del capítulo 5, página 497.

[8] Cfr. teorema 5.13, página 438, y la definición 5.1 del apartado 5.2.1.

[9] Cfr. teorema 5.18, página 447, y la definición 5.15 del apartado 5.2.3.

Supongamos que en el canal de transmisión se produjo un error y el mensaje que llega al receptor es

$$\vec{\omega}_C' = (1,1,0,1,0,0);$$

entonces, puesto que

$$[T(\vec{\omega}_C')] = H\vec{\omega}_C' = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

coincide (únicamente) con la tercera columna de la matriz verificadora de código H, el receptor puede identificar que el error se encuentra en la tercera coordenada del vector $\vec{\omega}_C' = (1,1,0,1,0,0)$, corregirlo al cambiar la tercera componente de este vector por el dígito 1 y obtener la palabra correcta que envió el emisor: $\vec{\omega} = (1,1,1)$. En caso contrario, si $\vec{\omega}_C'$ no contiene errores, $H\vec{\omega}_C = \vec{0}_{\mathbb{Z}_2^3}$ y el receptor, al notar esto, podrá determinar que el mensaje que recibió es correcto.◀

Aunque la manera de obtener matrices para generar y verificar códigos de paridad que explicamos arriba sirve para comprender y motivar estos conceptos, en realidad el proceso para lograr ese fin es mucho más simple basándonos en las mismas ideas. Recordemos que si H es una matriz, entonces la función $\vec{x} \mapsto T_H(\vec{x}) = H\vec{x}$ es una transformación lineal cuya representación matricial, relativa a las bases canónicas, es precisamente la matriz H. Entonces, para obtener un par de matrices, G y H, tales que la primera genere y la segunda verifique, respectivamente, un código de paridad (n,k) de Hamming, se requiere:

1. $G = [\ I_k \quad A\ ] \in \mathfrak{M}_{k\times n}(\mathbb{Z}_2)$, donde I_k es la matriz identidad de orden k y $A \in \mathfrak{M}_{k\times(n-k)}(\mathbb{Z}_2)$.
2. $H \in \mathfrak{M}_{(n-k)\times n}(\mathbb{Z}_2)$
3. El espacio nulo de la matriz H está generado por las filas de la matriz G (luego tiene dimensión k).
4. Las columnas de H son no nulas y distintas entre sí.

La construcción del par (G,H) está dada en el siguiente teorema:

Teorema B.12 ***(Código estándar de verificación de paridad de Hamming)*** *Sea A una matriz de tamaño $k\times(n-k)$ con componentes en $\mathbb{Z}_2$ tal que:*

1. *Las filas de A son distintas entres sí.*
2. *Las filas de A son no nulas.*
3. *Toda fila de A es distinta a cualquiera de las primeras $n-k$ columnas de la matriz identidad I_{n-k}.*

Entonces las matrices definidas por

$$G = [\ I_k \quad A\ ] \quad \text{y} \quad H = [\ A^t \quad I_{n-k}\ ]$$

forman, respectivamente, una matriz generadora y una matriz verificadora de un código de paridad (n,k) de Hamming; esto es:

- *El espacio nulo de H está generado por las filas de G; de hecho, forman una base.*
- *Las columnas de H son no nulas y distintas entre sí.*

DEMOSTRACIÓN ■ Notemos antes que los vectores fila de la matriz G son linealmente independientes porque G está en forma escalonada y en toda fila hay pivote. Ya que las primeras $n-k$ columnas de la matriz H son las filas de la matriz A y las últimas k columnas de H son las columnas de la matriz I_{n-k}, de las hipótesis 1, 2 y 3, se desprende que las columnas de H son no nulas y distintas entre sí. Por otra parte,

$$\begin{aligned} HG^t &= \begin{bmatrix} A^t & I_{n-k} \end{bmatrix} \begin{bmatrix} I_k \\ A^t \end{bmatrix} \\ &= 2A^t \\ &= \mathscr{O} \end{aligned}$$

lo cual implica que el subespacio generado por las k filas de G está contenido en el espacio nulo de H; pero, claramente el rango de esta matriz es $n-k$ y, por tanto, la dimensión del espacio nulo es k; luego las filas de G forman una base de ese subespacio por ser k vectores linealmente independientes. ■

Definición B.16 *Si A, G y H son matrices que satisfacen las condiciones del teorema precedente, se dice que G es una matriz estándar generadora de un código de paridad (n,k) de Hamming y que H es la matriz estándar de verificación de paridad correspondiente.*

En resumen, si G y H son una matriz generadora y su matriz de verificación de código de paridad (n,k) de Hamming, $\vec{\omega} = (x_1, x_2, \ldots, x_k)$ es una palabra binaria de longitud k, $\vec{\omega}_C = G^t\vec{\omega}$ es la codificación de verificación de paridad de $\vec{\omega}$ y $\vec{\omega}_C'$ es la palabra que recibe el receptor, entonces:

1. $\vec{\omega}_C'$ no contiene error (simple) si y sólo si $H\vec{\omega}_C' = \vec{0}_{\mathbb{Z}_2^{n-k}}$.
2. $\vec{\omega}_C'$ tiene un error (simple) si y sólo si $H\vec{\omega}_C' \neq \vec{0}_{\mathbb{Z}_2^{n-k}}$; en tal caso el error se encuentra en la coordenada j del vector $\vec{\omega}_C'$ de la única columna j de H que es idéntica a $H\vec{\omega}_C'$.

► **Ejemplo B.17** **(Código (7,4) de Hamming)** Sea

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

entonces A satisface las condiciones del teorema B.12, luego

$$G = \begin{bmatrix} I_4 & A \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

y

$$H = \begin{bmatrix} A^t & I_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

son, respectivamente, una matriz estándar generadora y su correspondiente matriz estándar de verificación de código de paridad $(7,3)$ de Hamming. Supongamos que a un receptor llegan los mensajes $\vec{\omega}_C' = (1,0,1,1,0,1,0)$ y $\vec{v}_C' = (1,0,1,1,0,0,1)$. ¿Qué mensajes realmente fueron enviados por la fuente?◀

Solución

(a)

$$H\vec{\omega}_C' = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix};$$

por tanto el mensaje que se recibió no tiene errores y el mensaje original es $\vec{\omega} = (1,0,1,1)$.

(b)

$$H\vec{v}_C' = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix};$$

que es igual a la primera columna de H. Entonces el mensaje $\vec{v}_C'$ contiene un error en la primera coordenada y, por ende, el mensaje original es $\vec{v} = (0,0,1,1)$. ✔

C Demostraciones que fueron diferidas

Este apartado contiene las demostraciones que se omitieron de los teoremas 2.7, 2.8, 2.9, 2.11, 3.14, 4.22, 4.24 y B.9 que se encuentran en las páginas 76, 80, 83, 84, 160, 327, 337 y 1025, respectivamente. Para facilitar la lectura se han repetido aquí los enunciados completos de cada teorema.

Teorema C.1 ***(2.7 Desarrollo por cofactores)*** *Sea $A \in \mathfrak{M}_{n\times n}$.*

1. *Sean $f_1, f_2, \ldots, f_n$ los elementos de una fila, F_i, cualquiera de A, con $c_1, c_2, \ldots, c_n$ sus respectivos cofactores y*

$$\Delta_{F_i} = c_1 f_1 + c_2 f_2 + \cdots + c_n f_n.$$

2. *Sean $g_1, g_2, \ldots, g_n$ los elementos de una columna, G_j, cualquiera de A, con $d_1, d_2, \ldots, d_n$ sus respectivos cofactores y*

$$\Delta_{G_j} = d_1 g_1 + d_2 g_2 + \cdots + d_n g_n.$$

Entonces:

- $\Delta_{F_i} = \Delta_{F_k}$ *para todo par de filas F_i y F_k de A.*
- $\Delta_{G_j} = \Delta_{G_l}$ *para todo par de columnas G_j y G_l de A.*
- $\Delta_{F_i} = \Delta_{G_j}$ *para toda fila F_i y toda columna G_j de A.*

DEMOSTRACIÓN ■ 1. Sean $A = [a_{ij}] \in \mathfrak{M}_{n\times n}$, $F_i = [\, a_{i1} \quad a_{i2} \quad \cdots \quad a_{in} \,]$ y $F_j = [\, a_{j1} \quad a_{j2} \quad \cdots \quad a_{jn} \,]$ un par de filas cualesquiera de A. Podemos suponer, sin perder generalidad, $i < j$. Si $n = 1$, el teorema es evidente y si $n = 2$, la afirmación se probó en el ejemplo 2.16 de la página 77. Supongamos que $n > 2$ es un entero y que el teorema es verdadero para toda matriz de orden menor o igual a $n-1$. Entonces:

(a)

$$\Delta_{F_i} = \sum_{k=1}^{n} a_{ik}(-1)^{i+k} M_{ik} \tag{C.1}$$

donde M_{ik} es el menor del elemento a_{ik} de la matriz A; es decir, el determinante de la matriz $A^{(ik)}$ que se obtiene a partir de la matriz A suprimiendo de ella la fila i y la columna k. Sea $k \in \{1, 2, \ldots, n\}$, entonces la matriz $A^{(ik)}$ tiene orden $n-1$ y podemos aplicar la hipótesis de inducción para calcular el menor M_{ik} desarrollando por cofactores con cualquiera de sus filas;

en particular, con la fila $j-1$, que en la matriz A corresponde, excepto por el elemento a_{jk}, a la fila j. De esta manera, ya que en la matriz $A^{(ik)}$ se ha suprimido la columna k,

$$\begin{aligned} a_{ik}(-1)^{i+k}M_{ik} &= a_{ik}(-1)^{i+k}\left(\sum_{l=1}^{k-1}a_{jl}(-1)^{j-1+l}M_{ikjl}+\sum_{l=k+1}^{n}a_{jl}(-1)^{j-1+l-1}M_{ikjl}\right) \\ &= a_{ik}(-1)^{i+k}\left(\sum_{l=1}^{k-1}a_{jl}(-1)^{j-1+l}M_{ikjl}-\sum_{l=k+1}^{n}a_{jl}(-1)^{j-1+l}M_{ikjl}\right) \end{aligned}$$

donde M_{ikjl} es el determinante de orden $n-2$ de la matriz que resulta de suprimir las filas i, j y las columnas k, l de la matriz A. De esta manera, el lado derecho de la relación (C.1) contiene $n(n-1)$ términos de la forma

$$a_{ik}a_{jl}(-1)^{i+j+k+l-1}M_{ikjl}, \text{ si } l<k \tag{C.2}$$

o

$$-a_{ik}a_{jl}(-1)^{i+j+k+l-1}M_{ikjl}, \text{ si } l>k \tag{C.3}$$

donde $k,l\in\{1,2,\ldots,n\}$, $l\neq k$.

(b)

$$\Delta_{F_j}=\sum_{l=1}^{n}a_{jl}(-1)^{j+l}M_{jl} \tag{C.4}$$

donde M_{jl} es el menor del elemento a_{jl} de la matriz A; esto es, el determinante de la matriz $A^{(jl)}$ que se obtiene a partir de la matriz A suprimiendo de ella la fila j y la columna l. Sea $l\in\{1,2,\ldots,n\}$, entonces la matriz $A^{(jl)}$ tiene orden $n-1$ y podemos aplicar la hipótesis de inducción para calcular el menor M_{jl} desarrollando por cofactores con cualquiera de sus filas; en particular, con la fila i que coincide, excepto por el elemento a_{il}, con la fila i de la matriz A, pues $j>i$. Luego, ya que en la matriz $A^{(jl)}$ se ha suprimido la columna l,

$$a_{jl}(-1)^{j+l}M_{jl}=a_{jl}(-1)^{j+l}\left(\sum_{k=1}^{l-1}a_{ik}(-1)^{i+k}M_{jlik}+\sum_{k=l+1}^{n}a_{ik}(-1)^{i+k-1}M_{jlik}\right)$$

donde M_{jlik} es el determinante de orden $n-2$ de la matriz que resulta de suprimir las filas i, j y las columnas l, k de la matriz A. Así, el miembro derecho de la relación (C.4) contiene $n(n-1)$ términos de la forma

$$-a_{jl}a_{ik}(-1)^{i+j+k+l-1}M_{jlik}, \text{ si } k<l \tag{C.5}$$

o

$$a_{jl}a_{ik}(-1)^{i+j+k+l-1}M_{jlik}, \text{ si } k>l \tag{C.6}$$

donde $k,l\in\{1,2,\ldots,n\}$, $l\neq k$.

De (C.2), (C.6) y (C.3), (C.5) se desprende

$$\Delta_{F_i} = \Delta_{F_j}.$$

Procediendo de manera análoga se demuestra que

$$\Delta_{G_k} = \Delta_{G_l}$$

para cualquier par de columnas.

Finalmente, veamos que el desarrollo por cofactores produce el mismo resultado por cualquier fila y por cualquier columna. Para ello, por lo que se acaba de probar, es suficiente demostrar que $\Delta_{F_1} = \Delta_{G_1}$; es decir, que al desarrollar por cofactores en la primera fila y en la primera columna se obtiene exactamente el mismo valor. Procederemos nuevamente por inducción sobre el orden de la matriz. Los casos $n = 1$ y $n = 2$ son evidentes. Sea $n > 2$ un entero y supongamos que la afirmación es cierta para toda matriz de orden menor o igual a $n - 1$. Si $A = [a_{ij}]$ tiene orden n, entonces

$$\Delta_{F_1} = a_{11}|M_{11}| + \sum_{j=2}^{n} a_{1j}(-1)^{1+j}M_{1j} \tag{C.7}$$

donde, para cada j, M_{1j} es el determinante de la matriz $A^{(1j)}$ que se obtiene de la matriz A al suprimir la primera fila y la columna j. Sea $j \in \{1,2,\ldots,n\}$, la matriz $A^{(1j)}$ tiene orden $n-1$, entonces, por la hipótesis de inducción, podemos calcular M_{1j} desarrollando por cofactores en la primera columna de la matriz $A^{(1j)}$; esto es,

$$\begin{aligned} M_{1k} &= \sum_{i=2}^{n} a_{i1}(-1)^{i-1+1}M_{1ji1} \\ &= \sum_{i=2}^{n} a_{i1}(-1)^{i}M_{1ji1} \end{aligned}$$

donde M_{1ji1} es el determinante de la matriz que resulta de suprimir las filas 1 e i y las columnas 1 y j de la matriz A. De esta manera, los términos en el lado derecho de (C.7) que contienen a $a_{1j}a_{i1}$ tienen la forma

$$a_{1j}a_{i1}(-1)^{j+i+1}M_{1ji1} \tag{C.8}$$

Por otra parte,

$$\Delta_{G_1} = a_{11}|M_{11}| + \sum_{i=2}^{n} a_{i1}(-1)^{i+1}M_{i1} \tag{C.9}$$

donde, para cada i, M_{i1} es el determinante de la matriz $A^{(i1)}$ que se obtiene de la matriz A al suprimir la fila i y la primera columna de ella. Sea $i \in \{1,2,\ldots,n\}$, la matriz $A^{(i1)}$ tiene orden $n-1$; al aplicar la hipótesis de inducción y desarrollar por cofactores a lo largo de su primera fila obtenemos

$$\begin{aligned} M_{i1} &= \sum_{j=2}^{n} a_{1j}(-1)^{j-1+1}M_{i11j} \\ &= \sum_{j=2}^{n} a_{1j}(-1)^{j}M_{i11j} \end{aligned}$$

donde M_{i11j} es el determinante de la matriz que se obtiene al suprimir las filas 1 e i y las columnas 1 y j de la matriz A. Entonces, los términos que contienen a $a_{i1}a_{1j}$ en el lado derecho de (C.9) son de la forma

$$a_{i1}a_{1j}(-1)^{i+j+1}M_{1ji1} \tag{C.10}$$

De (C.7), (C.8), (C.9) y (C.10) se desprende la igualdad $\Delta_{F_1} = \Delta_{G_1}$. Con lo cual hemos terminado la demostración del teorema. ■

Teorema C.2 ***(2.8 Propiedades de los determinantes)*** *Sean A y B matrices cuadradas de orden n, entonces:*

1. $|AB| = |A||B|$.
2. $|A^t| = |A|$.
3. *Si A es una matriz invertible, entonces* $|A^{-1}| \neq 0$ *y además* $|A^{-1}| = \dfrac{1}{|A|}$.
4. *Si A es triangular superior o triangular inferior, el determinante de A es el producto de los elementos de la diagonal.*
5. *Supongamos que B es equivalente a la matriz A al aplicarle una operación de renglón:*
   (a) *Si B se obtiene de A mediante la operación* $R_i \leftrightarrow R_j$, *entonces* $|A| = -|B|$.
   (b) *Si B resulta de A al aplicarle la operación* $R_i \leftrightarrow R_i + \alpha R_j$, *entonces* $|A| = |B|$.
   (c) *Si B se obtiene de A con la operación* $R_i \leftrightarrow \alpha R_i$, $\alpha \neq 0$, *entonces* $|A| = (1/\alpha)|B|$.

DEMOSTRACIÓN ■ Probaremos primero el inciso 5 del teorema y después los incisos 1, 2, 3 y 4 en ese orden. En el transcurso de la demostración se intercalarán, entre algunos incisos del teorema, ciertos lemas y sus respectivas demostraciones.

5. *(a)* Se procede por inducción sobre el orden n de la matriz $A = [a_{ij}]$. Si $n = 1$, la afirmación es evidente (si $n = 2$, un simple cálculo sirve para confirmar la afirmación). Sea $n > 2$ un entero y supongamos que la proposición es verdadera para matrices de orden menor a n. Entonces, si A es orden n, al desarrollar por cofactores con una fila p, distinta de i y j,

$$|B| = \sum_{k=1}^{n} a_{pk}(-1)^{p+k}N_{pk}$$

donde, para cada k, N_{pk} es el determinante de la matriz $B^{(pk)}$ que resulta de eliminar la fila p y la columna k de la matriz B. La matriz $B^{(pk)}$ es una matriz de orden $n-1$ que tiene intercambiadas las filas i y j de la matriz A (y no contiene al elemento a_{pk}), entonces, por hipótesis de inducción, $N_{pk} = -M_{pk}$, donde M_{pk} es el determinante de la matriz $A^{(pk)}$ que se obtiene al eliminar la fila p y la columna k de la matriz A. Luego,

$$\begin{aligned} |B| &= \sum_{k=1}^{n} a_{pk}(-1)^{p+k}N_{pk} \\ &= -\sum_{k=1}^{n} a_{pk}(-1)^{p+k}M_{pk} \\ &= -|A| \ . \end{aligned}$$

(c) Si B es la matriz que se obtiene mediante la operación de renglón $R_i \leftrightarrow \alpha R_i$, $\alpha \neq 0$, entonces

$$|B| = \sum_{k=1}^{n} \alpha a_{ik}(-1)^{i+k} M_{ik}$$

donde, para cada k, M_{ik} es el determinante de eliminar la fila i y la columna k de la matriz B y, por ende, es el menor del elemento a_{ik} de la matriz A; luego,

$$\begin{aligned} |B| &= \sum_{k=1}^{n} \alpha a_{ik}(-1)^{i+k} M_{ik} \\ &= \alpha \sum_{k=1}^{n} a_{ik}(-1)^{i+k} M_{ik} \\ &= \alpha\,|A|. \quad \blacksquare \end{aligned}$$

Lema C.1 *Sean $A = [a_{kj}]$ una matriz de orden n, $\vec{u} = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$ e $i \in \{1, 2, \ldots, n\}$ un índice fijo. Se denota con el símbolo $A_{\vec{u}}$ a la matriz que se obtiene al reemplazar la fila i por el vector $\vec{u}$; es decir, $A_{\vec{u}}$ tiene las mismas filas de la matriz A excepto que $a_{ik} = u_k$, para $k = 1, 2, \ldots, n$. Sea $T : \mathbb{R}^n \to \mathbb{R}$ la función definida por $T(\vec{u}) = \det(A_{\vec{u}})$. Entonces T es un funcional lineal; esto es,*

$$T(\alpha\vec{v} + \beta\vec{v}) = \alpha T(\vec{u}) + \beta T(\vec{v})\,.$$

$\forall \vec{u}, \vec{v} \in \mathbb{R}^n$, $\forall \alpha, \beta \in \mathbb{R}$.

DEMOSTRACIÓN ■ En efecto, si $\vec{u} = (u_1, u_2, \ldots, u_n)$ y $\vec{v} = (v_1, v_2, \ldots, v_n)$ son un par de vectores de $\mathbb{R}^n$ y α, β son dos escalares cualesquiera, entonces, al desarrollar $|A_{\alpha\vec{u}+\beta\vec{v}}|$ con los elementos de la fila i,

$$\begin{aligned} T(\alpha\vec{u} + \beta\vec{v}) &= \det(A_{\alpha\vec{u}+\beta\vec{v}}) \\ &= \sum_{k=1}^{n} (\alpha u_k + \beta v_k)(-1)^{i+k} M_{ik} \\ &= \alpha \sum_{k=1}^{n} u_k(-1)^{i+k} M_{ik} + \beta \sum_{k=1}^{n} v_k(-1)^{i+k} M_{ik} \\ &= \alpha\,|A_{\vec{u}}| + \beta\,|A_{\vec{v}}| \\ &= \alpha T(\vec{u}) + \beta T(\vec{v}). \quad \blacksquare \end{aligned}$$

Lema C.2 *Si una matriz de orden n, $A = [a_{ij}]$, tiene un par de filas iguales, digamos las filas i y j, entonces*

$$\det(A) = 0\,.$$

DEMOSTRACIÓN ■ Sea B la matriz que se obtiene al aplicar la operación de renglón $R_i \leftrightarrow R_j$ a la matriz A, entonces $|A| = |B|$ y, por la propiedad 5 *(a)*,

$$|A| = |B| = -|A|\,;$$

de donde se desprende la conclusión. ■

5. *(b)* Sea B la matriz que se obtiene de la matriz A al aplicar a ésta la operación de renglón $R_i \leftrightarrow R_i + \alpha R_j$. Entonces, por el lema C.1, si $\vec{u} = (a_{i1}, a_{i2}, \ldots, a_{in})$ y $\vec{v} = (a_{j1}, a_{j2}, \ldots, a_{jn})$,

$$\begin{aligned} |B| &= T(\vec{u} + \alpha\vec{v}) \\ &= T(\vec{u}) + \alpha T(\vec{v}) \\ &= |A_{\vec{u}}| + \alpha\, |A_{\vec{v}}| \\ &= |A| + \alpha 0 \\ &= |A| \end{aligned}$$

pues la matriz $A_{\vec{v}}$ tiene dos filas iguales (i y j) y, por el lema C.2, su determinante es nulo.

Lema C.3 *Sea $E \in \mathfrak{M}_{n\times n}$ una matriz elemental[1] y A una matriz del mismo tamaño, entonces:*

I. $|EA| = |E|\,|A|$.

II. *Si $A \in \mathfrak{M}_{n\times n}$, entonces A es no singular (invertible) si y sólo si $|A| \neq 0$.*

DEMOSTRACIÓN ■ I. Como E es una matriz elemental, E se obtiene de la identidad al aplicarle una de las operaciones de renglón *(i)* $R_i \leftrightarrow R_j$, *(ii)* $R_i \leftrightarrow \alpha R_i + \beta R_j$ ($\alpha \neq 0$); entonces, por el inciso 5, $|E| = -1$ o $|E| = \alpha \neq 0$. Sea $C = EA$, entonces C se obtiene de la matriz A al aplicarle la operación de renglón[2] *(i)* o *(ii)*; así, nuevamente por el inciso 5, $|C| = -|A|$ o $|C| = \alpha\,|A|$. Luego, en cualquier caso, se tiene $|EA| = |E|\,|A|$.
II. ($\Rightarrow$) Si A es no singular, entonces, por el teorema 2.6 (p. 73), existen k matrices elementales, $E_1, E_2, \ldots, E_k$, tales que

$$A = E_1 E_2 \cdots E_k\,;$$

así, por el inciso precedente,

$$|A| = |E_1|\,|E_2| \cdots |E_k| \neq 0.$$

($\Leftarrow$) Si $|A| \neq 0$ y A es singular (no invertible), entonces A es equivalente por filas a una matriz H que está en forma escalonada, con una fila nula y que tiene, por tanto, determinante cero, al desarrollar éste por cofactores en esa fila. Luego, existen matrices elementales, $E_1, E_2, \ldots, E_m$, tales que $E_1 E_2 \cdots E_m A = H$ y, por el primer inciso de este lema,

$$0 \neq |E_1|\,|E_2| \cdots |E_m|\,|A| = |H| = 0$$

lo cual es una contradición. Por ende, A debe ser una matriz no singular. ■

1. Sean $A, B \in \mathfrak{M}_{n\times n}$, si A es singular, entonces AB también es singular[3] y se tiene, por el lema anterior $|A| = 0$, $|AB| = 0$; en consecuencia $|AB| = |A|\,|B|$. Supongamos ahora que A es una matriz

[1] Cfr. definición 2.3, página 71.
[2] Cfr. teorema 2.4, página 71.
[3] Pues en caso contrario $C = AB$ sería una matriz no singular, lo que implicaría $(C^{-1}A)B = I$ y, por tanto, la matriz $C^{-1}A = D$ sería una matriz invertible y en consecuencia A sería no singular.

no singular, entonces existen matrices elementales,[4] $E_1, E_2, \ldots, E_k$, tales que $A = E_1E_2\cdots E_k$; luego, por el primer inciso del lema precedente,

$$|A| = |E_1|\,|E_2|\cdots|E_k|$$

y

$$\begin{aligned} |AB| &= |E_1E_2\cdots E_kB| \\ &= |E_1|\,|E_2|\cdots|E_k|\,|B| \\ &= |A|\,|B|\,. \end{aligned}$$

2. Se procede nuevamente por inducción. El resultado es evidente si $n = 1$. Sea $n > 1$ un entero y supongamos que la proposición es verdadera para todo determinante de orden $n-1$. Sea $B = A^t$ y desarrollemos $|A^t|$ por la primera columna y $|A| = |a_{ij}|$ por la primera fila, entonces

$$|A^t| = \sum_{k=1}^{n} a_{1k}(-1)^{1+k}N_{k1}$$

donde, para cada k, N_{k1} es el determinante de la matriz $B^{(1k)}$ que se obtiene suprimiendo la primera fila y la columna k de la matriz $A^t = B$. Si $A^{(1k)}$ es la matriz que se obtiene al suprimir la primera fila y la columna k de la matriz A, entonces es claro que $B^{(1k)} = \left(A^{(1k)}\right)^t$. Luego, por la hipótesis de inducción,

$$\begin{aligned} N_{k1} &= |B^{(1k)}| \\ &= \left|\left(A^{(1k)}\right)^t\right| \\ &= |A^{(1k)}| \\ &= M_{1k}\,, \end{aligned}$$

el cofactor del elemento a_{1k} en la matriz A. Por tanto,

$$\begin{aligned} |A^t| &= \sum_{k=1}^{n} a_{1k}(-1)^{1+k}N_{k1} \\ &= \sum_{k=1}^{n} a_{1k}(-1)^{1+k}M_{1k} \\ &= |A|\,. \end{aligned}$$

3. Cfr. el ejemplo 2.21 (página 81).

4. Se procede por inducción. Si $n = 1$, la afirmación es evidente. Supongamos que $n > 1$ es un entero y que la proposición es verdadera para toda matriz triangular superior de orden $n-1$. Al desarrollar por cofactores el determinante de $A = [a_{ij}] \in \mathfrak{M}_{n\times n}$ por medio de la primera columna obtenemos

[4] Cfr. teorema 2.6, página 73.

$$|A| = a_{11}M_{11};$$

donde M_{11} es el determinante de la matriz $A^{(11)}$ que resulta al suprimir la primera fila y la primera columna de la matriz A. Puesto que A es una matriz triangular superior, $A^{(11)}$ también es una matriz triangular superior y los elementos en su diagonal son $a_{22}, a_{33}, \ldots, a_{nn}$. Entonces, por la hipótesis de inducción,

$$\begin{aligned} |A| &= a_{11}M_{11} \\ &= a_{11}|A^{(11)}| \\ &= a_{11}a_{22}\cdots a_{nn} \end{aligned}$$

es decir, el producto de los elementos en la diagonal de A. ■

Teorema C.3 ***(2.9 Método de la adjunta)*** *Si* $A \in \mathfrak{M}_{n\times n}$, *entonces*

$$A\,\text{Adj}(A) = \det(A)I_n.$$

Luego, A es invertible si y sólo si $\det(A) \neq 0$ *y en tal caso*

$$A^{-1} = \tfrac{1}{|A|}\text{Adj}(A).$$

DEMOSTRACIÓN ■ Sea $A = [a_{ij}]$, entonces $\text{Adj}(A) = [c_{ji}]$, donde c_{ji} es el cofactor del elemento a_{ji}; luego, si $A\text{Adj}(A) = [m_{ij}]$, se tiene

$$\begin{aligned} m_{ij} &= \begin{bmatrix} a_{i1} & a_{i_2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} c_{j1} \\ c_{j2} \\ \vdots \\ c_{jn} \end{bmatrix} \\ &= a_{i1}c_{j1} + a_{i2}c_{j2} + \cdots + a_{in}c_{jn}. \end{aligned}$$

Si $i = j$, ya que $a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + a_{in}c_{in}$ es el desarrollo del determinante de A por cofactores en la fila i,

$$\begin{aligned} m_{ii} &= a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + a_{in}c_{in} \\ &= \det(A). \end{aligned}$$

Si $i \neq j$, sea A' la matriz que se obtiene de la matriz A al reemplazar la fila j por la fila i; esto es, A' es la matriz A salvo que las filas i y j son iguales a la fila i de A, entonces

$$\begin{aligned} m_{ij} &= a_{i1}c_{j1} + a_{i2}c_{j2} + \cdots + a_{in}c_{jn} \\ &= a_{j1}c_{j1} + a_{j2}c_{j2} + \cdots + a_{jn}c_{jn} \\ &= \det(A'). \end{aligned}$$

Al aplicar la operación de renglón $R_j \leftrightarrow R_j - R_i$ a la matriz A', se obtiene una matriz A'' con la fila j nula, cuyo determinante, al desarrollar por cofactores en esta fila, resulta ser cero;[5] luego, por la propiedad 5 *(a)* del teorema 2.8,

$$m_{ij} = \det(A') = -\det(A'') = 0.$$

Por tanto

$$\begin{aligned} A\operatorname{Adj}(A) &= \operatorname{diag}(\det(A), \det(A), \ldots, \det(A)) \\ &= \det(A) I_n \\ &= |A|\, I_n\,. \quad \blacksquare \end{aligned}$$

Teorema C.4 ***(2.11 Regla de Cramer)*** *Sea A una matriz cuadrada de orden n. Sea, para cada* $i = 1, 2, \ldots, n$, Δ_i *el determinante que resulta de la matriz que se obtiene de A al sustituir la i-ésima columna por* $\vec{b}$. *Entonces, si* $\Delta = |A| \neq 0$ *el sistema* $A\vec{x} = \vec{b}$ *tiene solución única*

$$x_i = \frac{\Delta_i}{\Delta},$$

para $i = 1, 2, \ldots, n$.

DEMOSTRACIÓN ■ Puesto que $|\Delta| \neq 0$, por el teorema anterior, $A = [a_{kj}]$ es invertible y

$$\begin{aligned} A^{-1} &= \frac{1}{\Delta}\operatorname{Adj}(A) \\ &= \frac{1}{\Delta}[c_{jk}] \end{aligned}$$

donde c_{jk} es el cofactor del elemento a_{jk} de la matriz A. Entonces, si $\vec{b} = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}$,

$$\begin{aligned} \vec{x} &= A^{-1}\vec{b} \\ &= \frac{1}{\Delta}[c_{jk}] \begin{bmatrix} b_1 \\ b_2 \\ \\ b_n \end{bmatrix}. \end{aligned}$$

Luego,

$$x_i = \frac{1}{\Delta}(b_1 c_{1i} + b_2 c_{2i} + \cdots + b_n c_{ni})\,;$$

y ya que $c_{1i}, c_{2i}, \ldots, c_{ni}$ son los cofactores de la columna i de la matriz A,

$$x_i = \frac{\Delta_i}{\Delta}$$

para cada $i = 1, 2, \ldots, n$. ■

[5] Cfr. lema C.2 en la demostración de la propiedad 5 *(b)* del teorema C.2 en este apéndice.

Antes de probar el teorema 3.14 necesitamos establecer algunos preliminares:

1. En un conjunto $\mathscr{A}$ una relación entre sus elementos, denotada por[6] $\preceq$, es un **orden parcial** en él si satisface las siguientes tres propiedades:

   (a) (*Reflexividad*) $a \preceq a \; \forall a \in \mathscr{A}$.

   (b) (*Antisimetría*) $a, b \in \mathscr{A}$, $a \preceq b$, $b \preceq a \Rightarrow a = b$.

   (c) (*Transitividad*) $a, b \in \mathscr{A}$, $a \preceq b$, $b \preceq c \Rightarrow a \preceq c$.

2. Si en el conjunto $\mathscr{A}$ existe una relación de orden parcial $\preceq$, se dice que $\mathscr{A}$ (o con más precisión el par $(\mathscr{A}, \preceq)$) es un conjunto *parcialmente ordenado*.

3. Si en un conjunto $\mathscr{A}$ se tiene una relación de orden parcial $\preceq$ que cumple, además, la condición

$$a \preceq b \text{ o } b \preceq a \; \forall a, b \in \mathscr{A}$$

   se dice que la relación $\preceq$ es de orden *total* y que $\mathscr{A}$ (o con más precisión el par $(\mathscr{A}, \preceq)$) es un conjunto **totalmente ordenado.**

4. Sea $\mathscr{F}$ una colección no vacía de subconjuntos de un conjunto X. Si $A, B \in \mathscr{F}$, la relación $A \preceq B \Leftrightarrow A \subset B$; es decir, $A \preceq B \Leftrightarrow A$ es subconjunto de B, es una relación de orden parcial (pero no necesariamente de orden total).

5. Sea $(\mathscr{A}, \preceq)$ un conjunto parcialmente ordenado y $\mathscr{T} \subset \mathscr{A}$. Si $\mathscr{T}$ con la relación de orden parcial $\preceq$ es un conjunto totalmente ordenado, se dice que $\mathscr{T}$ es una **cadena** en $\mathscr{A}$.

6. Sea $\mathscr{B}$ un subconjunto de un conjunto parcialmente ordenado $(\mathscr{A}, \preceq)$. $c \in \mathscr{A}$ es una **cota superior** de $\mathscr{B}$, si $b \preceq c$ para todo $b \in \mathscr{B}$. Si $\mathscr{B}$ posee una cota superior en $\mathscr{A}$, se dice que $\mathscr{B}$ está **acotado superiormente**.

7. Sea $(\mathscr{A}, \preceq)$ un conjunto parcialmente ordenado. Un elemento $\mathfrak{m}$ en $\mathscr{A}$ es un **elemento maximal** de $\mathscr{A}$, si para cualquier $x \in \mathscr{A}$ es válida la siguiente implicación:

$$\mathfrak{m} \preceq x \Rightarrow \mathfrak{m} = x.$$

8. Si $\mathbf{E}$ es un espacio vectorial y $C \subset \mathbf{E}$, $\mathscr{L}(C)$, la intersección de todos los subespacios de $\mathbf{E}$ que contienen a C, es un subespacio de $\mathbf{E}$, el subespacio **generado por** C. Si $C \neq \emptyset$, $\mathscr{L}(C)$ es el conjunto de todas las posibles combinaciones lineales de elementos de C. Si $\mathbf{E} = \mathscr{L}(C)$, se dice que C genera a $\mathbf{E}$ (cfr. ejercicios 31 y 32 del capítulo 3, pág. 190).

La demostración del teorema 3.14 requiere del llamado lema de Zorn. Este lema es una de las más importantes herramientas en las matemáticas para probar la existencia de ciertos tipos de entes. Para demostrar el lema de Zorn se requiere del axioma de elección, que es una piedra angular de las matemáticas. A continuación establecemos sin probar este resultado y remitimos al lector interesado en consultar una demostración de este importantísimo teorema y los conceptos preliminares 1 a 7, a la referencia bibliográfica [14].

[6] $a \preceq b$ se lee "a precede o es igual a b".

Lema C.4 ***(de Zorn)*** *Sea $\mathscr{A}$ un conjunto no vacío y $\preceq$ una relación de orden parcial en él. Si toda cadena en $\mathscr{A}$ está acotada superiormente, entonces existe al menos un elemento maximal de $\mathscr{A}$.*

Teorema C.5 ***(3.14 Existencia de bases)*** *Sea* **E** *un espacio vectorial.*[7] *Entonces* **E** *posee una base; es decir, existe $\mathscr{B} \subset \mathbf{E}$ L.I. tal que todo vector del espacio* **E** *es combinación lineal de elementos de $\mathscr{B}$.*

DEMOSTRACIÓN ■ Sea **E** un espacio vectorial. Si $\mathbf{E} = \{\vec{0}_{\mathbf{E}}\}$, entonces $\mathscr{B} = \emptyset$ es una base de[8] **E**. Supongamos que $\mathbf{E} \neq \{\vec{0}_{\mathbf{E}}\}$, por tanto existe $\vec{u}_0 \in \mathbf{E} - \{\vec{0}_{\mathbf{E}}\}$. Sea

$$\mathscr{F} = \{S \subset \mathbf{E} \,|\, S \text{ es linealmente independiente}\}$$

con el orden parcial de la inclusión del inciso 4 de los preliminares; puesto que $\{\vec{u}_0\}$ es L.I., $\mathscr{F} \neq \emptyset$. Sea $\mathscr{T} = \{S_\lambda\}_{\lambda \in \Lambda}$ una cadena en $\mathscr{F}$; es decir, una colección de subconjuntos L.I., S_λ, $\lambda \in \Lambda$, que está totalmente ordenada con la relación de inclusión de conjuntos. Sea

$$S = \bigcup_{\lambda \in \Lambda} S_\lambda \,;$$

claramente $S_\lambda \subset S$ para todo $\lambda \in \Lambda$. Supongamos que S es L.D., entonces existen $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n \in S$ y escalares $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$, no todos nulos, que satisfacen

$$\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \cdots + \alpha_n \vec{u}_n = \vec{0}_{\mathbf{E}}$$

Sean $\lambda_i \in \Lambda$, $i = 1, 2, \ldots, n$, tales que $\vec{u}_i \in S_{\lambda_i}$ para cada índice i. Entonces, como $\mathscr{T}$ es una cadena, existe $i_0 \in \{1, 2, \ldots, n\}$ que satisface $S_{\lambda_i} \subset S_{\lambda_{i_0}}$ para todo i. Luego, los vectores $\vec{u}_i \in S_{\lambda_{i_0}}$ para todo i y forman un conjunto L.D. contenido en el conjunto L.I. $S_{\lambda_{i_0}}$; lo cual es una contradicción. Por tanto, $S \in \mathscr{F}$ y, entonces, es una cota superior para la cadena $\mathscr{T}$. Por el lema de Zorn, el conjunto parcialmente ordenado $\mathscr{F}$ posee un elemento maximal $\mathscr{B}$. Este conjunto, por pertenecer a $\mathscr{F}$, es L.I. Supongamos que $\vec{u} \in \mathbf{E}$ y que $\vec{u} \notin \mathscr{L}(\mathscr{B})$, entonces el conjunto $\mathscr{B} \cup \{\vec{u}\}$ es linealmente independiente, por tanto pertenece a $\mathscr{F}$, y

$$\mathscr{B} \subsetneq \mathscr{B} \cup \{\vec{u}\}$$

lo cual es imposible pues $\mathscr{B}$ es un elemento maximal de $\mathscr{F}$; por ende $\vec{u}$ debe pertenecer a $\mathscr{L}(\mathscr{B})$. Ya que $\mathscr{B}$ es linealmente independiente y genera el espacio **E**, $\mathscr{B}$ es una base para el espacio **E**. ■

Necesitamos, como antes, de algunos preliminares para probar los últimos tres teoremas de este apéndice. Sea **E** un espacio vectorial real y $\|\cdot\|$ una norma en él.

1. $G \subset \mathbf{E}$ es un conjunto **abierto**, si para cada $\vec{u} \in G$ existe $r > 0$ tal que la bola $B(\vec{u}, r) \subset G$ (cfr. definición 4.13, página 311).

[7] El espacio puede estar sobre cualquier campo de escalares.

[8] $\emptyset$ es L.I, por vacuidad (en caso contrario existirían vectores L.D. que pertenecen $\emptyset$) y, por el problema resuelto 32 *(a)* del capítulo 3 (pág. 190), $\mathscr{L}(\emptyset) = \{\vec{0}_{\mathbf{E}}\}$.

2. $F \subset \mathbf{E}$ es un conjunto **cerrado** si su complemento F^c, relativo a $\mathbf{E}$, es un conjunto abierto.
3. $A \subset \mathbf{E}$ es un conjunto **acotado**, si existe $m > 0$ tal que $\|\vec{u}\| \leq m$ para todo $\vec{u} \in A$. En tal caso, se dice que m es una cota para el conjunto A.
4. Si $\mathbf{E}$ tiene dimensión finita, $K \subset \mathbf{E}$ es un conjunto **compacto** si K es cerrado y acotado.
5. Si $f : \mathbf{E} \to \mathbb{R}$ es una función continua[9] en todos los puntos de un conjunto compacto $K \subset \mathbf{E}$, entonces f alcanza su máximo y su mínimo en K. Es decir, existen $\vec{x}_1, \vec{x}_2 \in K$ tales que $f(\vec{x}_1) \leq f(\vec{x}) \leq f(\vec{x}_2)$ para todo $\vec{x} \in K$.

La proposición contenida en el inciso precedente es un teorema fundamental del análisis matemático; el lector interesado en conocer los detalles de la demostración de este teorema puede consultar la página 142 de la referencia bibliográfica [21].

Teorema C.6 ***(4.22 Equivalencia de normas en dimensión finita)*** *Si* $\mathbf{E}$ *es un espacio vectorial, entonces cualquier par de normas en* $\mathbf{E}$ *son equivalentes; es decir, en un espacio de dimensión finita todas las normas en él son equivalentes.*

DEMOSTRACIÓN ■ Sean $\|\cdot\|$ una norma en $\mathbf{E}$ y $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ una base de este espacio. Es claro que la transformación $T : \mathbf{E} \to \mathbb{R}^n$ definida, para cada $\vec{u} = \sum_{i=1}^n \alpha_i \vec{e}_i$, como $T(\vec{u}) = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ es un isomorfismo;[10] además, si se define $\mathscr{N} : \mathbb{R}^n \to \mathbb{R}$ como $\mathscr{N}(\alpha_1, \alpha_2, \ldots, \alpha_n) = \|\sum_{i=1}^n \alpha_i \vec{e}_i\|$, entonces $\mathscr{N}$ es una norma en $\mathbb{R}^n$. Por consiguiente T es una isometría; esto es, $\mathscr{N}(T(\vec{u})) = \|\vec{u}\|$; por lo que basta probar que todas las normas en $\mathbb{R}^n$ son equivalentes y, dado que la relación de equivalencia entre normas es transitiva, es suficiente mostrar que toda norma $\|\cdot\|$ en $\mathbb{R}^n$ es equivalente a la norma euclidiana $\|(\alpha_1, \alpha_2, \ldots, \alpha_n)\|_e = \sqrt{\alpha_1^2 + \cdots + \alpha_n^2}$. Sea $\|\cdot\|$ una norma en $\mathbb{R}^n$ y sea la función $\varphi : \mathbb{R}^n \to \mathbb{R}$ definida como $\varphi(\vec{u}) = \|\vec{u}\|$, para cada $\vec{u} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$. Entonces, si $\vec{u} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, $\vec{v} = (\beta_1, \beta_2, \ldots, \beta_n)$ son elementos arbitrarios de $\mathbb{R}^n$ y $\vec{e}_i = \underbrace{(0, 0, \ldots, 0, \overset{i}{1}, 0, \ldots, 0)}_{n}$, $i = 1, 2, \ldots, n$, son los vectores unitarios de la base canónica de $\mathbb{R}^n$, se tiene

$$
\begin{aligned}
|\varphi(\vec{u}) - \varphi(\vec{v})| &= |\|\vec{u}\| - \|\vec{v}\|| \\
&\leq \|\vec{u} - \vec{v}\| \\
&\leq \|\sum_{i=1}^n (\alpha_i - \beta_i)\vec{e}_i\| \\
&\leq \sum_{i=1}^n |\alpha_i - \beta_i| \, \|\vec{e}_i\| \\
&\leq \sum_{i=1}^n \|\vec{u} - \vec{v}\|_\infty \|\vec{e}_i\| \\
&= \left(\sum_{i=1}^n \|\vec{e}_i\|\right) \|\vec{u} - \vec{v}\|_\infty \\
&\leq \left(\sum_{i=1}^n \|\vec{e}_i\|\right) \|\vec{u} - \vec{v}\|_e .
\end{aligned}
$$

[9] Cfr. definición 6.9.

[10] Cfr. definición 5.16, página 452.

Por lo que la función φ es continua en $(\mathbb{R}^n, \|\cdot\|_e)$ —de hecho, uniformemente continua—. Sea $S = \{\vec{u} \in \mathbb{R}^n \mid \|\vec{u}\| = 1\}$; es fácil ver[11] que S es un conjunto cerrado y acotado, luego S es un conjunto compacto en el espacio $(\mathbb{R}^n, \|\cdot\|_e)$, por lo que la función φ alcanza un mínimo α y un máximo β en S. Sea $\vec{x} \in \mathbb{R}^n - \{\vec{0}_{\mathbb{R}^n}\}$, entonces $\frac{1}{\|\vec{x}\|_e}\vec{x} \in S$ y, por tanto,

$$\alpha \leq \varphi\left(\frac{1}{\|\vec{x}\|_e}\vec{x}\right) \leq \beta;$$

esto es,

$$\alpha \leq \left\|\frac{1}{\|\vec{x}\|_e}\vec{x}\right\| \leq \beta.$$

En consecuencia

$$\alpha\,\|\vec{x}\|_e \leq \|\vec{x}\| \leq \beta\,\|\vec{x}\|_e$$

para todo $\vec{x} \in \mathbb{R}^n$. ■

Lema C.5 *Si $(\mathbf{E}, \|\cdot\|)$ es un espacio vectorial normado de dimensión finita, entonces la bola cerrada de centro $\vec{0}_{\mathbf{E}}$ y radio r, $B[\vec{0}_{\mathbf{E}}, r]$ (cfr. definición 4.13, pág. 311) es un conjunto compacto.*

DEMOSTRACIÓN ■ Sea $\vec{u}_0 \in (B[\vec{0}_{\mathbf{E}}, r])^C = \mathbb{R}^n - B[\vec{0}_{\mathbf{E}}, r]$, entonces $\|\vec{u}_0\| > r$. Sea $\delta = \|\vec{u}_0\| - r$; luego, si $\vec{u} \in B(\vec{u}_0, \delta)$,

$$\|\vec{u} - \vec{u}_0\| < \delta = \|\vec{u}_0\| - r\,.$$

Por tanto,

$$|\|\vec{u}\| - \|\vec{u}_0\|| \leq \|\vec{u} - \vec{u}_0\| < \|\vec{u}_0\| - r$$

implica

$$r - \|\vec{u}_0\| < \|\vec{u}\| - \|\vec{u}_0\| < \|\vec{u}_0\| - r\,;$$

de donde

$$\|\vec{u}\| > r.$$

Es decir, $\vec{u} \in (B[\vec{0}_{\mathbf{E}}, r])^C$ y, por ende, este conjunto es abierto; lo cual prueba que $B[\vec{0}_{\mathbf{E}}, r]$ es cerrado. Claramente $B[\vec{0}_{\mathbf{E}}, r]$ es un conjunto acotado ($\|\vec{x}\| \leq r\ \forall \vec{x} \in B[\vec{0}_{\mathbf{E}}, r]$). Ya que $\mathbf{E}$ tiene dimensión finita y $B[\vec{0}_{\mathbf{E}}, r]$ es cerrado y acotado, se concluye que este conjunto es compacto. ■

[11] La prueba de este hecho es similar a la demostración del lema C.5.

Teorema C.7 ***(4.24 Aproximaciones óptimas en espacios normados)*** *Sean* $\mathbf{E}$ *un espacio vectorial normado con norma* $\|\cdot\|$*,* S *un subespacio de dimensión finita en* $\mathbf{E}$ *y* $\vec{u}$ *un vector dado de* $\mathbf{E}$*. Entonces existe una aproximación óptima* $\vec{p}^{\,*}$ *de* $\vec{u}$ *en* S*.*

DEMOSTRACIÓN ■ $\vec{0}_{\mathbf{E}} \in S$, por tanto es un candidato a ser $\vec{p}^{\,*}$. De esta manera, si $\vec{x} \in S$ y $\|\vec{x}\| > 2\|\vec{u}\|$, se tiene

$$\begin{aligned} \|\vec{u}-\vec{x}\| &\geq \|\vec{x}\| - \|\vec{u}\| \\ &> 2\|\vec{u}\| - \|\vec{u}\| \\ &> \|\vec{u}\| \\ &= \|\vec{u}-\vec{0}_{\mathbf{E}}\|. \end{aligned}$$

Por tanto, si $\vec{u}$ tiene una aproximación óptima, ésta se debe encontrar en la bola cerrada $B[\vec{0}_{\mathbf{E}}, 2\|\vec{u}\|]$ en el espacio S. Basta entonces limitar la búsqueda a este conjunto que, por el lema anterior, es un conjunto compacto. Sea $\varphi : S \to \mathbb{R}$ la función definida por $\varphi(\vec{x}) = \|\vec{u}-\vec{x}\|$, entonces, ya que para todo par $\vec{x}_1, \vec{x}_2 \in S$

$$\begin{aligned} |\varphi(\vec{x}_1) - \varphi(\vec{x}_2)| &= \|\vec{u}-\vec{x}_1-\vec{u}+\vec{x}_2\| \\ &= \|\vec{x}_2-\vec{x}_1\|\ , \end{aligned}$$

se desprende que φ es una función continua en todo punto del subespacio S. Luego, por el lema precedente, existe $\vec{p}^{\,*} \in B[\vec{0}_{\mathbf{E}}, 2\|\vec{u}\|]$ tal que $\varphi(\vec{x})$ es mínima en $\vec{p}^{\,*}$ y, por tanto, en S; es decir,

$$\|\vec{u}-\vec{p}^{\,*}\| \leq \|\vec{u}-\vec{x}\|$$

para todo $\vec{x} \in S$. ■

Remitimos al lector a las definiciones 6.8 y 6.9 y al teorema 6.9 del apartado 6.8.2 del capítulo 6, para los conceptos de límite y continuidad que se requieren en lo que sigue.

Lema C.6 *Sean* $\mathbb{C}$ *el campo complejo considerado como el espacio normado* $\mathbb{R}^2$ *con la norma canónica* $\|\cdot\|$*,* $n \in \mathbb{N}$*, y* $A, a_j \in \mathbb{C}$*,* $j = 0,1,2,\ldots,n$*, números dados.*

1. *Si* $\varphi : \mathbb{C} = \mathbb{R}^2 \to \mathbb{R}$ *es la función definida por* $\varphi(z) = Az^n$*, entonces* φ *es una función continua en todo* $z \in \mathbb{C}$*.*
2. *Sea* $f : \mathbb{C} \to \mathbb{R}$*, la función definida* $f(z) = a_0 + a_1 z + \cdots + a_n z^n$*, entonces* f *es continua en todo* $z \in \mathbb{C}$*.*

DEMOSTRACIÓN ■ 1. Claramente el resultado es cierto si $n = 1$. Sea n un entero mayor a 1 y supongamos que la afirmación es verdadera para todo natural $k < n$. Sea $z_0 \in \mathbb{C}$ fijo, pero arbitrario, entonces

$$\begin{aligned} \|\varphi(z) - \varphi(z_0)\| &= |Az^n - Az_0^n| \\ &= |A|\left|(z-z_0)(z^{n-1} + z^{n-2}z_0 + \cdots + zz_0^{n-2} + z_0^{n-1}\right| \\ &= |A|\,|z-z_0|\left|z^{n-1} + z^{n-2}z_0 + \cdots + zz_0^{n-2} + z_0^{n-1}\right| \\ &\leq |A|\,|z-z_0|\left(\left|z^{n-1}\right| + \left|z^{n-2}z_0\right| + \cdots + \left|zz_0^{n-2}\right| + \left|z_0^{n-1}\right|\right) \\ &= |A|\,|z-z_0|\left(\left|z^{n-1}\right| + \left|z^{n-2}\right||z_0| + \cdots + \left|zz_0^{n-2}\right|\left|z_0^{n-2}\right| + \left|z_0^{n-1}\right|\right). \end{aligned}$$

Luego, por la hipótesis de inducción,

$$\begin{aligned}\lim_{z\to z_0}\|\varphi(z)-\varphi(z_0)\| &\le \lim_{z\to z_0}|A|\,|z-z_0|\sum_{j=1}^{n}\left|z^{n-j}z_0^{j-1}\right|\\ &= |A|\,n\left|z_0^{n-1}\right|\lim_{z\to z_0}|z-z_0|\\ &= 0.\end{aligned}$$

2. Es consecuencia inmediata del primer inciso por el hecho de que la suma de funciones continuas es también una función continua. ■

Definición C.1 *Si $f:\mathbb{R}^n\to\mathbb{R}$, se escribe*

$$\lim_{\|\vec{u}\|\to\infty} f(\vec{u})=\infty,$$

si dado $M>0$ existe $r>0$ tal que $\|\vec{u}\|>r\Rightarrow f(\vec{u})>M$.

Lema C.7 *Si $z\to f(z)$ es un polinomio de grado mayor o igual a uno, entonces*

$$\lim_{|z|\to\infty}|f(z)|=\infty.$$

DEMOSTRACIÓN ■ Ya que el grado de f es mayor o igual a uno, f tiene la forma $f(z)=a_0+a_1z+\cdots+a_nz^n$, con n un entero mayor o igual a 1 y $a_n\neq 0$. Entonces

$$|f(z)|=|a_nz^n|\left|\frac{a_0}{a_nz^n}+\frac{a_1}{a_nz^{n-1}}+\cdots+\frac{a_{n-1}}{a_nz}+1\right|.$$

Es fácil mostrar que $\lim_{|z|\to\infty}\left|\frac{a}{z^k}\right|=0$ y que $\lim_{|z|\to\infty}|az^k|=\infty$ para todo $a\in\mathbb{C}$ y para todo $k\in\mathbb{N}$. Luego

$$\lim_{|z|\to\infty}|f(z)|=\infty. \quad ■$$

Lema C.8 *Si $z=x+iy\in\mathbb{C}$ está en la circunferencia de centro 0 y radio 1; esto es, $|z|=1$, entonces existe un único $\beta\in[0,2\pi)$ tal que $z=e^{i\beta}$.*

DEMOSTRACIÓN ■ Supongamos que z está en el primer cuadrante, entonces, como $|z|=1$, $x^2+y^2=1$ y, por tanto, $0\le x\le 1$, $0\le y\le 1$. Entonces, dado que la función $t\mapsto\cos(t)$ es continua en todo punto y la imagen del intervalo $[0,\pi/2]$ bajo esta función es el intervalo $[0,1]$, por el teorema del valor intermedio[12] existe

[12] Si una función $f:[a,b]\to\mathbb{R}$ es continua y $f(a)\le y\le f(b)$, entonces existe $\mu\in[a,b]$ tal que $f(\mu)=y$.

$\beta \in [0, \pi/2]$ tal que $\cos(\beta) = x$; de la inyectividad de la función $t \mapsto \cos(t)$ en el intervalo $[0, \pi/2]$ se desprende la unicidad de β. Puesto que $\cos^2(\beta) + \operatorname{sen}^2(\beta) = 1 = x^2 + y^2$, se desprende que $y = \operatorname{sen}(\beta)$; por tanto $e^{i\beta} = z$. Los casos en los demás cuadrantes se deducen de lo precedente utilizando reflexiones sobre cada uno de los ejes de coordenadas. ■

Teorema C.8 ***(B.9 Teorema fundamental del álgebra)*** *Sea $p(z)$ un polinomio con coeficientes reales y/o complejos (i.e., $p(z) \in \mathbb{C}[z]$) que tiene grado mayor o igual a uno, entonces $p(z)$ tiene por lo menos una raíz en $\mathbb{C}$.*

DEMOSTRACIÓN ■ Como el grado de $p(z)$ es mayor o igual a 1, p tiene la forma

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$$

con los coeficientes $a_j \in \mathbb{C}$, $n \in \mathbb{N}$ y $a_n \neq 0$. Por el lema C.7 existe $r_0 > 0$ tal que

$$|p(z)| > |a_0|$$

para todo z con $|z| > r_0$. Puesto que la composición de funciones continuas es una función continua y la función $z \mapsto \|z\| = |z|$ es continua en todo punto, el lema C.6 implica que la función $z \mapsto |p(z)|$ es una función continua en todo punto. Por el lema C.5, página 1049, $D[\vec{0}_{\mathbb{R}_2}, r_0]$ es un conjunto compacto, por tanto la función $z \mapsto |p(z)|$ alcanza un valor mínimo absoluto[13] en un punto z_0 de este disco; esto es, $|p(z_0)| \leq |p(z)|$ para todo $z \in D[\vec{0}_{\mathbb{R}_2}, r_0]$. Ya que $p(\vec{0}_{\mathbb{R}^2}) = p(0) = a_0$, y $\vec{0}_{\mathbb{R}^2} \in D[\vec{0}_{\mathbb{R}_2}, r]$, se tiene $|p(z_0)| \leq |a_0|$. Y como $|p(z)| > |a_0|$ si $|z| > r_0$, se desprende que

$$|p(z_0)| \leq |p(z)| \quad \forall z \in \mathbb{C} \tag{C.11}$$

es decir, $|p(z_0)|$ es un mínimo absoluto de la función $z \mapsto |p(z)|$ en el plano $\mathbb{R}^2 = \mathbb{C}$. Afirmamos que $p(z_0) = 0$. Supongamos que no es así. Sea el polinomio $g(z) = \dfrac{p(z+z_0)}{p(z_0)}$, entonces $\operatorname{grad}(g) = n$, $g(0) = 1$ y, por (C.11),

$$|g(z)| = \frac{|p(z+z_0)|}{|p(z_0)|} \geq 1 \quad \forall z \in \mathbb{C} \tag{C.12}$$

Sea $k \geq 1$ el subíndice del primer coeficiente de $g(z) = 1 + \sum_{k=1}^{n} c_k z^k$ que es distinto de cero, el cual existe porque $\operatorname{grad}(p) \geq 1$, entonces g tiene la forma

$$g(z) = 1 + c_k z^k + c_{k+1} z^{k+1} + \cdots + c_n z^n$$

con $c_k \neq 0$. Sea $w = -|c_k| c_k^{-1}$, entonces $|w| = 1$, en consecuencia w está en la circunferencia de centro 0 y radio 1; luego, por el lema C.8, existe θ tal que $w = e^{ik\theta}$; y, por ende, $-|c_k| = c_k e^{ik\theta}$. Así, para cualquier número real $r > 0$ la fórmula de De Moivre implica

[13] Cfr. inciso 5 de los hechos preliminares de la página 1048.

$$\begin{aligned} g(re^{i\theta}) &= 1 + c_k r^k e^{ik\theta} + c_{k+1} r^{k+1} e^{i(k+1)\theta} + \cdots + c_n r^n e^{in\theta} \\ &= 1 - |c_k|\, r^k + c_{k+1} r^{k+1} e^{i(k+1)\theta} + \cdots + c_n r^n e^{in\theta}. \end{aligned}$$

Por la desigualdad triangular se tiene

$$\begin{aligned} \left|g(re^{i\theta})\right| &\le \left|1 - |c_k|\, r^k\right| + \left|c_{k+1} r^{k+1} e^{i(k+1)\theta}\right| + \cdots + \left|c_n r^n e^{in\theta}\right| \\ &= \left|1 - |c_k|\, r^k\right| + |c_{k+1}|\, r^{k+1} + \cdots + |c_n|\, r^n \end{aligned}$$

Como $r^k \to 0$ cuando $r \to 0$, podemos elegir r suficientemente pequeño de tal suerte que $1 - |c_k|\, r^k > 0$; por tanto

$$\begin{aligned} \left|g(re^{i\theta})\right| &\le 1 - |c_k|\, r^k + |c_{k+1}|\, r^{k+1} + \cdots + |c_n|\, r^n \\ &= 1 - r^k \left(|c_k| - |c_{k+1}|\, r - \cdots - |c_n|\, r^{n-k}\right) \end{aligned}$$

Ahora, podemos elegir r aún más pequeño, si es necesario, para que $|c_k| - |c_{k+1}|\, r - \cdots - |c_n|\, r^{n-k} > 0$ y, por ende,

$$\left|g(re^{i\theta})\right| \le 1 - r^k \left(|c_k| - |c_{k+1}|\, r - \cdots - |c_n|\, r^{n-k}\right) < 1$$

lo cual es una contradicción a (C.12). ■

D | Formas canónicas de Jordan

En todo lo que resta de este apéndice los espacios que se consideren serán espacios vectoriales sobre un campo fijo $\mathbb{K}$ y las matrices que aparezcan tendrán componentes en ese campo. El lector que no esté habituado a trabajar con campos generales debería estudiar primero el apéndice B. Denotaremos en todo este apartado la identidad multiplicativa del campo simplemente por 1 en lugar de $1_{\mathbb{K}}$ y el neutro aditivo del campo, $0_{\mathbb{K}}$, por 0.

Si T es un operador lineal en un espacio vectorial de dimensión finita, T no necesariamente es diagonalizable; sin embargo, bajo condiciones muy generales, siempre es posible encontrar una base del espacio de tal suerte que la representación matricial de T, relativa a esa base, tenga un formato muy simple y fácil de manejar. Los formatos que estudiaremos son las llamadas formas canónicas de Jordan.

Definición D.1 ***(Bloques de Jordan)*** *Una matriz cuadrada de orden k de la forma*

$$J_k(\mu) = \begin{bmatrix} \mu & 1 & 0 & 0 & \cdots & 0 \\ 0 & \mu & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \cdots & \mu & 1 & 0 \\ 0 & 0 & \cdots & 0 & \mu & 1 \\ 0 & 0 & \cdots & 0 & 0 & \mu \end{bmatrix},$$

*donde $\mu \in \mathbb{K}$ (las componentes en la diagonal son todas iguales a μ, las componentes en la supradiagonal son todas iguales a la identidad multiplicativa y las demás componentes son nulas), se llama **bloque de Jordan** de orden k.*

Si $\mu = 0$ en un bloque de Jordan, éste tiene la forma

$$N_k = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{D.1}$$

esto es, si $N_k = [n_{ij}]$, $n_{i(i+1)} = 1$, $i = 1, 2, \ldots, k-1$, y las demás componentes son nulas. Notemos que para todo bloque de Jordan se tiene

$$J_k(\mu) = \mu I_k + N_k$$

y, en particular, $J_1(\mu) = [\mu]$ y $N_1 = [0]$.

Si $p(x) = \sum_{k=0}^{m} a_k x^k$ es un polinomio y T es un operador lineal en un espacio vectorial, se define el operador lineal, evaluación[1] de p en T, como el operador lineal $p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m$, donde $T^k = \underbrace{T \circ T \circ \cdots \circ T}_{k}$, la composición de T consigo mismo k veces.[2] Y si A es una matriz cuadrada, se define la matriz $p(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_m A^m$, la evaluación de p en la matriz A.

Definición D.2
1. *Sean $\mathbf{E}$ un espacio vectorial y T un operador lineal en $\mathbf{E}$. Se dice que T es **nilpotente** si existe $m \in \mathbb{N}$ tal que $T^m = \theta$, el operador constante cero.*
2. *Sea A una matriz cuadrada, A es **nilpotente** si existe $m \in \mathbb{N}$ tal que $A^m = \mathscr{O}$.*

*Si T (A) es un operador (matriz) nilpotente, al menor entero $k \in \mathbb{N}$ tal que $T^k = \theta$ ($A^k = \mathscr{O}$) se le llama **índice de nilpotencia** del operador[3] (de la matriz).*

Es claro que el polinomio mínimo (cfr. ejercicio resuelto del capítulo 5, pág. 536) de la matriz (D.1) es $\mathfrak{m}(x) = x^k$; luego $(N_k)^k = \mathfrak{m}(N_k) = 0$ y, si $(N_k)^q = 0$, entonces N_k es un cero del polinomio $f(x) = x^q$; por tanto $\mathfrak{m}$ divide a f y en consecuencia $k \leq q$; luego, N_k es nilpotente con índice de nilpotencia k. Las matrices N_k definidas por (D.1) se llaman **matrices nilpotentes básicas**.

► **Ejemplo D.1** Las matrices

$$J_3(2) = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} =, \quad J_2(-1) = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} \quad \text{y} \quad J_1 = [4]$$

son bloques de Jordan de órdenes $k = 3$, $k = 2$ y $k = 1$, respectivamente, y, por ejemplo,

$$J_3(2) = 2I_2 + N_2 = 2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}. ◄$$

Definición D.3 *Una matriz cuadrada A es **diagonal por bloques**, si existen $A_1, A_2, \ldots, A_r$ matrices cuadradas tales*[4]

$$A = \begin{bmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{r-1} & 0 \\ 0 & 0 & \cdots & 0 & A_r \end{bmatrix},$$

donde cada matriz A_k tiene su diagonal en la diagonal de la matriz A. En tal caso se dice que la matriz A es la suma directa de las matrices A_k y se escribe $A = A_1 \oplus A_2 \oplus \cdots \oplus A_r$.

[1] Cfr. instrucciones de los ejercicios resueltos 49 a 53 del capítulo 5, página 534.

[2] Cfr. ejercicio resuelto 12, página 504.

[3] Cfr. ejercicio propuesto 168, página 551.

[4] Los ceros por arriba y por abajo de las matrices A_k significan que por arriba y por abajo de cada una de estas matrices las componentes de A son nulas.

▶ **Ejemplo D.2** La matriz

$$\left[\begin{array}{cc:c:cc} 1 & 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 3 & 5 \end{array}\right]$$

es diagonal por bloques con $A_1 = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$, $A_2 = [1]$, $A_3 = \begin{bmatrix} 4 & 1 \\ 3 & 5 \end{bmatrix}$.◀

Definición D.4

1. *Sea $J = A_1 \oplus A_2 \oplus \cdots \oplus A_r$ una matriz diagonal por bloques; si cada uno de los bloques A_j es un bloque de Jordan, se dice que J es una **forma canónica de Jordan.***
2. *Sea $N = A_1 \oplus A_2 \oplus \cdots \oplus A_r$ una matriz diagonal por bloques; si cada uno de los bloques A_j es una matriz nilpotente básica, se dice que N es una **forma canónica nilpotente**.*

No es difícil probar que toda forma canónica nilpotente es, en efecto, una matriz nilpotente; la demostración de este hecho se deja de ejercicio al lector.

▶ **Ejemplo D.3** Sean

$$J = \left[\begin{array}{cc:c:cc} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{array}\right] \text{ y } N = \left[\begin{array}{cc:ccc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Entonces J es una forma canónica de Jordan y N es una forma canónica nilpotente.◀

Si $\mathbf{E}$ es un espacio vectorial, $\mathbf{E}$ es la **suma directa**[5] de los subespacios S_k, $k = 1, 2, \ldots, r$, y se escribe

$$\mathbf{E} = S_1 \oplus S_2 \oplus \cdots \oplus S_r\,,$$

si todo vector $\vec{u} \in \mathbf{E}$ puede expresarse de manera unívoca como

$$\vec{u} = \vec{u}_1 + \vec{u}_2 + \cdots + \vec{u}_r$$

donde cada $\vec{u}_k \in S_k$ para $k = 1, 2, \ldots, r$. Además, en el caso de ser $\mathbf{E}$ un espacio finitamente generado, si

$$\mathscr{B}_k = \{\vec{v}_{k1}, \vec{v}_{k2}, \ldots, \vec{v}_{kn_k}\}, \tag{D.2}$$

$k = 1, 2, \ldots, r$, son bases de sendos subespacios S_k de $\mathbf{E}$, entonces $\mathbf{E}$ es la suma directa de estos subespacios si sólo si $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2 \cup \cdots \cup \mathscr{B}_r$, i. e.,

[5] Cfr. ejercicio resuelto 27, página 187.

$$\mathscr{B} = \{\vec{v}_{11}, \vec{v}_{12}, \ldots, \vec{v}_{1n_1}, \vec{v}_{21}, \vec{v}_{22}, \ldots, \vec{v}_{2n_2}, \ldots, \vec{v}_{r1}, \vec{v}_{r2}, \ldots, \vec{v}_{rn_r}\} \tag{D.3}$$

es una base de **E**.

Por otra parte, si T es un operador lineal en un espacio **E**, un subespacio $S \subset \mathbf{E}$ es T-**invariante**[6] si $T(S) \subset S$; entonces la restricción del operador T al subespacio S —esto es, $\widehat{T}_S : S \to S$, con $\widehat{T}_S(\vec{u}) = T(\vec{u})$— es un operador lineal en él. Tenemos el siguiente teorema para representar matricialmente a un operador lineal sobre un espacio, que es suma directa de subespacios T-invariantes, por medio de una matriz diagonal por bloques. La demostración es muy sencilla y se deja de ejercicio al lector.

Teorema D.1 *Sea $T : \mathbf{E} \to \mathbf{E}$ un operador lineal en un espacio de dimensión finita* **E**. *Se supone que* **E** *es suma directa de r subespacios T-invariantes S_k. Sean $\mathscr{B}_k$, $k = 1, 2, \ldots, r$, las bases de sendos subespacios dadas por la relación (D.2) y A_k la representación matricial de la restricción de T al subespacio S_k, entonces la representación matricial $[T]$, relativa a la base $\mathscr{B}$ definida por la relación (D.3), es la matriz diagonal por bloques*

$$[T] = \begin{bmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{r-1} & 0 \\ 0 & 0 & \cdots & 0 & A_r \end{bmatrix}.$$

❍ **Nota D.1** Sean **E**, T, los subespacios T-invariantes S_k y las matrices A_k como en el teorema D.1; si T_k es la restricción del operador T al subespacio S_k, $k = 1, 2, \ldots, r$, se acostumbra decir que el operador T se descompone en los operadores T_k o que es la suma directa de estos operadores, y se escribe $T = T_1 \oplus T_2 \oplus \cdots \oplus T_k = \bigoplus_{k=1}^{r} T_k$.

Recordemos que si T y R son operadores lineales en un mismo espacio vectorial, la composición de ellos, que resulta también ser lineal, se representa con la notación TR en lugar de $T \circ R$; es decir, $TR = T \circ R$. Además, si T es como antes, con frecuencia se escribe $T\vec{u}$ en lugar de $T(\vec{u})$. Es fácil probar las siguientes propiedades, su demostración se deja de ejercicio al lector.

Teorema D.2 *Sean f y g un par de polinomios y T, R operadores lineales en un espacio vectorial* **E**, *entonces:*

1. $(f+g)(T) = f(T) + g(T)$.
2. $(fg)(T) = f(T)g(T)$.
3. $(fg)(T) = g(T)f(T)$.

De aquí en adelante utilizaremos libremente las propiedades básicas de los polinomios sobre un campo $\mathbb{K}$; el lector que requiera repasar éstas puede consultar el apartado B.3.1 del apéndice B.3, páginas 1021 a 1025.

[6] Cfr. ejercicio resuelto 47, página 532.

Lema D.1 *Si f es un polinomio*[7] *y T es un operador lineal en un espacio $\mathbf{E}$ de dimensión finita:*

1. $\text{Ker}(f(T))$ *es un subespacio T-invariante.*
2. *Sean S un subespacio T-invariante, $\mathfrak{m}$ y $\mathfrak{m}_1$ los polinomios mínimos de T y de la restricción T_1 del operador lineal T a S, respectivamente, entonces $\mathfrak{m}_1$ divide a $\mathfrak{m}$.*
3. *Sean S_1 y S_2 dos subespacios T-invariantes tales que $\mathbf{E} = S_1 \oplus S_2$, y sean T_1, T_2 las restricciones del operador T a los subespacios S_1 y S_2, respectivamente. Si $\mathfrak{m}$, $\mathfrak{m}_1$ y $\mathfrak{m}_2$ son los polinomios mínimos de sendos operadores T, T_1 y T_2, entonces $\mathfrak{m}$ es el mínimo común múltiplo de $\mathfrak{m}_1$ y $\mathfrak{m}_2$.*

DEMOSTRACIÓN ■ 1. Sea $\vec{u} \in \text{Ker}(f(T))$, entonces, ya que $xf(x) = f(x)x$, se tiene

$$\begin{aligned} f(T)T\vec{u} &= Tf(T)\vec{u} \\ &= T(\vec{0}_{\mathbf{E}}) \\ &= \vec{0}_{\mathbf{E}} \end{aligned}$$

y, por tanto, $T(\vec{u}) \in \text{Ker}(f(T))$.

2. Para todo $\vec{u} \in S$,

$$\mathfrak{m}(T_1)\vec{u} = \mathfrak{m}(T)\vec{u} = \vec{0}_{\mathbf{E}}$$

por tanto $\mathfrak{m}_1$ divide a $\mathfrak{m}$.

3. Por el inciso precedente $\mathfrak{m}_1$ y $\mathfrak{m}_2$ dividen a $\mathfrak{m}$. Sea f un polinomio tal que $\mathfrak{m}_1$ y $\mathfrak{m}_2$ dividen a f, entonces $f(T_1)\vec{x}_1 = \vec{0}_{\mathbf{E}} = f(T_2)\vec{x}_2$ para todo $\vec{x}_1 \in S_1$ y para todo $\vec{x}_2 \in S_2$, luego, si $\vec{u} = \vec{u}_1 + \vec{u}_2 \in \mathbf{E}$, $\vec{u}_1 \in S_1$ y $\vec{u}_2 \in S_2$, es cualquier vector de $\mathbf{E}$, se tiene

$$\begin{aligned} f(T)\vec{u} &= f(T)(\vec{u}_1 + \vec{u}_2) \\ &= f(T)\vec{u}_1 + f(T)\vec{u}_2 \\ &= f(T_1)\vec{u}_1 + f(T_2)\vec{u}_2 \\ &= \vec{0}_{\mathbf{E}} + \vec{0}_{\mathbf{E}} \\ &= \vec{0}_{\mathbf{E}} . \end{aligned}$$

Por tanto $\mathfrak{m}$ divide a f; luego, $\mathfrak{m}$ es el mínimo común múltiplo de $\mathfrak{m}_1$ y $\mathfrak{m}_2$. ■

Teorema D.3 *Sean $\mathbf{E}$ un espacio vectorial de dimensión finita y $T : \mathbf{E} \to \mathbf{E}$ un operador lineal. Si un polinomio $p(x) = g(x)h(x)$, donde los polinomios g y h son primos relativos entre sí, satisface $p(T) = 0$, entonces*

$$\mathbf{E} = \text{Ker}(g(T)) \oplus \text{Ker}(h(T))$$

y $\text{Ker}(g(T))$, $\text{Ker}(h(T))$ son subespacios T-invariantes. En particular, si $p = \mathfrak{m}$ es el polinomio mínimo de T y g, h son mónicos (el coeficiente de la mayor potencia es uno), entonces g es el polinomio mínimo de la restricción de T a $\text{Ker}(g(T))$ y h es el polinomio mínimo de la restricción de T a $\text{Ker}(h(T))$.

[7] Cfr. ejercicio resuelto 52 del capítulo 5, página 536.

DEMOSTRACIÓN ■ Por el primer inciso del lema D.1, $\operatorname{Ker}(g(T))$ y $\operatorname{Ker}(h(T))$ son subespacios T-invariantes. Dado que g y h son polinomios primos relativos, existen dos polinomios, p_1 y p_2, tales que $p_1(x)g(x)+p_2(x)h(x)=1$; luego

$$(p_1g)(T)+(p_2h)(T)=I,$$

donde I es el operador identidad en $\mathbf{E}$ ($I(\vec{u})=\vec{u}$). Entonces, para todo $\vec{u}\in\mathbf{E}$,

$$\vec{u}=(p_1g)(T)\vec{u}+(p_2h)(T)\vec{u} \tag{D.4}$$

Por otra parte,

$$\begin{aligned} h(T)((p_1g)(T)\vec{u}) &= h(T)(gp_1)T\vec{u} \\ &= h(T)((g(T)p_1(T)\vec{u}) \\ &= h(T)g(T)(p_1(T)\vec{u}) \\ &= (hg)(T)(p_1(T)\vec{u}) \\ &= p(T)(p_1(T)\vec{u}) \\ &= \vec{0}_{\mathbf{E}}. \end{aligned}$$

Por tanto, $(p_1g)(T)\vec{u}\in\operatorname{Ker}(h(T))$; análogamente se demuestra que $(p_2h)(T)\vec{u}\in\operatorname{Ker}(g(T))$. Con lo cual hemos probado $\mathbf{E}=\operatorname{Ker}(g(T))+\operatorname{Ker}(h(T))$. Si $\vec{w}\in\operatorname{Ker}(g(T))\cap\operatorname{Ker}(h(T))$, entonces, por la relación D.4, se tiene

$$\begin{aligned} \vec{w} &= (p_1g)(T)\vec{w}+(p_2h)(T)\vec{w} \\ &= p_1(T)(g(T)\vec{w})+p_2(T)(h(T)\vec{w}) \\ &= p_1(T)\vec{0}_{\mathbf{E}}+p_2(T)\vec{0}_{\mathbf{E}} \\ &= \vec{0}_{\mathbf{E}}; \end{aligned}$$

luego, $\operatorname{Ker}(g(T))\cap\operatorname{Ker}(h(T))=\{\vec{0}_{\mathbf{E}}\}$ y, por ende, $\mathbf{E}=\operatorname{Ker}(g(T))\oplus\operatorname{Ker}(h(T)$. Sean $\mathfrak{m}$, $\mathfrak{m}_1$ y $\mathfrak{m}_2$ los polinomios mínimos de los operadores lineales T, T_1 y T_2, respectivamente, donde T_1 y T_2 son las restricciones de T a sendos espacios $\operatorname{Ker}(g(T))$ y $\operatorname{Ker}(h(T))$. Por el tercer inciso del lema D.1, $\mathfrak{m}$ es el mínimo común múltiplo de $\mathfrak{m}_1$ y $\mathfrak{m}_2$. Ya que $g(T)\vec{u}=\vec{0}_{\mathbf{E}}$ para todo $\vec{u}\in\operatorname{Ker}(g(T))$ y $h(T)\vec{v}=\vec{0}_{\mathbf{E}}$ para todo $\vec{v}\in\operatorname{Ker}(h(T))$, $g(T_1)=\theta=h(T_2)$; luego $\mathfrak{m}_1$ divide a g y $\mathfrak{m}_2$ divide a h; entonces, por ser g y h primos relativos, $\mathfrak{m}_1$ y $\mathfrak{m}_2$ son también primos relativos; por ende $\mathfrak{m}=\mathfrak{m}_1\mathfrak{m}_2$. Así que

$$\mathfrak{m}_1\mathfrak{m}_2=gh$$

y, puesto que $\mathfrak{m}_1$ divide a g, existe un polinomio q tal que $g=q\mathfrak{m}_1$. Entonces

$$\mathfrak{m}_1\mathfrak{m}_2=q\mathfrak{m}_1h$$

implica

$$\mathfrak{m}_1(\mathfrak{m}_2-qh)=\theta$$

de donde se desprende que h divide a $\mathfrak{m}_2$; y ya que los dos polinomios son mónicos, se desprende $h=\mathfrak{m}_2$. De manera análoga se prueba $g=\mathfrak{m}_1$. ■

Teorema D.4 ***(Teorema de descomposición primaria)*** *Sean* $\mathbf{E}$ *un espacio vectorial de dimensión finita y* T *un operador lineal en él. Si el polinomio mínimo de* T *tiene la forma*

$$\mathfrak{m}_T = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$$

donde los p_k *son polinomios mónicos, irreducibles y distintos entre sí, entonces*

$$\mathbf{E} = S_1 \oplus S_2 \oplus \cdots \oplus S_r$$

con $S_j = \operatorname{Ker}(p_j^{n_j}(T))$*; además, el polinomio mínimo de la restricción del operador* T *al subespacio* T*-invariante* S_k *es* $p_j^{n_j}$*, para cada* $j = 1, 2, \ldots, r$.

DEMOSTRACIÓN ■ Procedamos por inducción sobre r. Si $r = 1$, entonces $\mathfrak{m}_T = p_1^{n_1}$ implica $p_1^{n_1}(T) \equiv 0$ y, por tanto, $\mathbf{E} = \operatorname{Ker}(p_1^{n_1}(T))$. Sea $k > 1$ un entero y supongamos que la afirmación es cierta para $r = k$. Supongamos que el polinomio característico de T tiene la forma $\mathfrak{m}_T = p_1^{n_1} q$, donde $q = p_2^{n_2} \cdots p_k^{n_k} p_{k+1}^{n_{k+1}}$. Puesto que los polinomios $p_j^{n_j}$ son mónicos, irreducibles y distintos entre sí, entonces $p_1^{n_1}$ y q son polinomios mónicos y primos relativos; por el teorema D.3

$$\mathbf{E} = S_1 \oplus W,$$

donde los subespacios T-invariantes que forman la suma directa son $S_1 = \operatorname{Ker}(p_1^{n_1}(T))$ y $W = \operatorname{Ker}(q(T))$; y, además, el polinomio mínimo de la restricción de T a S_1 es $p_1^{n_1}$ y el polinomio mínimo de la restricción de T a W es q. Sea $R = \widehat{T}_W$, la restricción de T a W. Entonces R es un operador lineal en el espacio W con polinomio mínimo $q = p_2^{n_2} \cdots p_k^{n_k} p_{k+1}^{n_{k+1}}$, con los polinomios $p_j^{n_j}$ mónicos, irreducibles y distintos entre sí; por hipótesis de inducción

$$W = S_2 \oplus \cdots \oplus S_k \oplus S_{k+1}$$

donde, para cada $j = 2, \ldots, k, k+1$, $S_j = \operatorname{Ker}(p_j^{n_j}(R))$ y el polinomio característico de la restricción de R a S_j es $p_j^{n_j}$. Ya que $p_j^{n_j}$ divide a q, para cada $j = 2, \ldots, k, k+1$, se desprende que $\operatorname{Ker}(p_j^{n_j}(T)) \subset W$ y, por tanto, $\operatorname{Ker}(p_j^{n_j}(T)) = \operatorname{Ker}(p_j^{n_j}(R))$ para todo $j = 2, \ldots k, k+1$. Claramente la restricción de R a cada subespacio S_j coincide con la restricción de T a este subespacio que es precisamente $\operatorname{Ker}(p_j^{n_j}(T))$; por tanto, el polinomio mínimo de la restricción de T al subespacio S_j es $p_j^{n_j}$. De esta manera $\mathbf{E} = S_1 \oplus S_2 \oplus \cdots \oplus S_{k+1}$ con cada $S_j = \operatorname{Ker}(p_j^{n_j}(T))$ y $p_j^{n_j}$ el polinomio mínimo de la restricción de T a S_j. Por inducción la afirmación del teorema es verdadera para todo r. ■

Corolario D.1

1. *Sea* T *un operador lineal en un espacio vectorial* $\mathbf{E}$ *de dimensión finita* n. T *es diagonalizable si y sólo si el polinomio mínimo de* T *es el producto de factores lineales distintos entre sí:*

$$\mathfrak{m}_T(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_r) \tag{D.5}$$

2. *Sea* A *una matriz cuadrada.* A *es diagonalizable si y sólo si el polinomio mínimo de* A *es producto de factores lineales distintos entre sí:*

$$\mathfrak{m}_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_r).$$

DEMOSTRACIÓN ■ 1. ($\Rightarrow$) Supongamos que T es diagonalizable y sean λ_j, $j = 1,2,\ldots,r$, los valores propios, distintos entre sí, de T. Luego $\mathbf{E}$ tiene una base, $\mathscr{B} = \{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\}$, de vectores propios correspondientes a los valores propios λ_j. Si definimos

$$p(T) = (T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_r I),$$

entonces $p(T)\vec{e}_j = \vec{0}_{\mathbf{E}}$ para todo $j = 1,2,\ldots,n$, y, por tanto, $p(T)\vec{u} = \vec{0}_{\mathbf{E}}$ para todo $\vec{u} \in \mathbf{E}$. Así, $\mathfrak{m}_T$ —el polinomio mínimo de T— divide al polinomio

$$p(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_r);$$

en consecuencia $\mathfrak{m}_T$ es producto de factores lineales.

($\Leftarrow$) Supongamos ahora que $\mathfrak{m}_T$, el polinomio mínimo de T, es producto de factores lineales como en (D.5). Por el teorema de descomposición primaria (teorema D.4), $\mathbf{E}$ es la suma directa de los subespacios $S_j = \operatorname{Ker}(p_j(T))$, donde $p_j(\lambda) = (\lambda - \lambda_j)$. Si $\{\vec{u}_{j1},\ldots,\vec{u}_{jn_j}\}$ es una base de S_j para $j = 1,\ldots,r$, entonces cada uno de los elementos de S_j es un vector propio correspondiente a λ_j; y como $\mathbf{E}$ es la suma directa de estos subespacios, la reunión de todas estas bases es una base para $\mathbf{E}$; así, $\mathbf{E}$ tiene una base de vectores propios de T y, por tanto, T es diagonalizable.

2. Es consecuencia inmediata del primer inciso al aplicar éste al operador $T_A : \mathbb{K}^n \to \mathbb{K}^n$, definido por $T_A(\vec{u}) = A\vec{u}$ para cada $\vec{u} \in \mathbf{E}$. ■

Lema D.2 *Sean* $\mathbf{E}$ *un espacio vectorial de dimensión finita, T un operador lineal en él, $k \geq 2$ un entero, $U = \operatorname{Ker}(T^{k-2})$, $V = Ker(T^{k-1})$ y $W = \operatorname{Ker}(T^k)$. Por el ejercicio resuelto 23 del capítulo 5 (pág. 511) $U \subset V \subset W$. Sea $\{\vec{u}_1,\ldots,\vec{u}_r\}$ una base de U; completemos[8] ésta a una base $\{\vec{u}_1,\ldots,\vec{u}_r,\vec{v}_1,\ldots,\vec{v}_s\}$ de V; y completemos esta última a una base $\{\vec{u}_1,\ldots,\vec{u}_r,\vec{v}_1,\ldots,\vec{v}_s,\vec{w}_1,\ldots,\vec{w}_t\}$ de W. Sea*

$$B = \{\vec{u}_1,\ldots,\vec{u}_r,T(\vec{w}_1),\ldots,T(\vec{w}_t)\}$$

entonces:

1. *$B \subset V$.*
2. *B es linealmente independiente.*

DEMOSTRACIÓN ■ 1. Puesto que $U \subset V$, únicamente hay que demostrar que los $T(\vec{w}_j) \in V$. En efecto, puesto que $\vec{w}_j \in W$, $T^k(\vec{w}_j) = \vec{0}_{\mathbf{E}}$; por tanto

$$T^{k-1}(T(\vec{w}_j)) = T^k(\vec{w}_j) = \vec{0}_{\mathbf{E}}.$$

Lo cual implica $T(\vec{w}_j) \in V$.

2. Sea $\alpha_1,\ldots,\alpha_r,\beta_1,\ldots,\beta_t \in \mathbb{K}$ tales que

$$\alpha_1\vec{u}_1 + \cdots + \alpha_r\vec{u}_r + \beta_1 T(\vec{w}_1) + \cdots + \beta_t T(\vec{w}_t) = \vec{0}_{\mathbf{E}} \tag{D.6}$$

[8] Cfr. proceso de compleción de un conjunto L.I. a una base, página 167.

Entonces
$$\begin{aligned}\vec{0}_{\mathbf{E}} &= T^{k-2}(\alpha_1\vec{u}_1+\cdots+\alpha_r\vec{u}_r+\beta_1 T(\vec{w}_1)+\cdots+\beta_t T(\vec{w}_t))\\ &= \sum_{j=1}^{r}\alpha_j T^{k-2}(\vec{u}_j)+\sum_{j=1}^{t}\beta_j T^{k-1}(\vec{w}_j)\\ &= \vec{0}_{\mathbf{E}}+\sum_{j=1}^{t}\beta_j T^{k-1}(\vec{w}_j)\\ &= \sum_{j=1}^{t}\beta_j T^{k-1}(\vec{w}_j)\\ &= T^{k-1}\left(\sum_{j=1}^{t}\beta_j\vec{w}_j\right)\end{aligned}$$

implica $\beta_1\vec{w}_1+\cdots+\beta_t\vec{w}_t\in V$; Por tanto, existen $\gamma_1,\ldots,\gamma_r,\delta_1,\ldots,\delta_s\in\mathbb{K}$ tales que

$$\gamma_1\vec{u}_1+\cdots+\gamma_r\vec{u}_r+\delta_1\vec{v}_1+\cdots+\delta_s\vec{v}_s+\beta_1\vec{w}_1+\cdots+\beta_t\vec{w}_t=\vec{0}_{\mathbf{E}}.$$

Luego, como $\{\vec{u}_1,\ldots,\vec{u}_r,,\vec{v}_1,\ldots,\vec{v}_s,\vec{w}_1,\ldots,\vec{w}_t\}$ es linealmente independiente, se desprende $\beta_j=0$ para todo j. Al sustituir los valores $\beta_j=0$ en (D.6) se obtiene $\alpha_1\vec{u}_1+\cdots+\alpha_r\vec{u}_r=\vec{0}_{\mathbf{E}}$; finalmente, puesto que los vectores $\vec{u}_j$ son linealmente independientes, los α_j deben ser también nulos. ■

❍ **Nota D.2** Observemos que en la demostración de la inclusión en V e independencia lineal de los vectores $T(\vec{w}_1),\ldots,T(\vec{w}_t)$ se puede prescindir de la hipótesis de la independencia lineal y pertenencia al espacio U de los vectores $\vec{u}_j$; es decir, si V, W y los vectores $\vec{v}_j$ y $\vec{w}_k$ son como antes, entonces los vectores $T(\vec{w}_k)$ pertenecen a V y son linealmente independientes.

A continuación establecemos y demostramos el teorema para representar matricialmente un operador nilpotente por medio de una matriz diagonal por bloques básicos nilpotentes; es decir, una forma canónica nilpotente.

Teorema D.5 ***(Formas canónicas nilpotentes)***

1. *Sean $\mathbf{E}$ un espacio vectorial de dimensión finita n y $T:\mathbf{E}\to\mathbf{E}$ un operador lineal nilpotente con índice de nilpotencia ν. Sean $S_k=\operatorname{Ker}(T^k)$ y $m_k=\dim(S_k)$, $k=1,2,\ldots,\nu$. Entonces existe una base $\mathscr{B}$ de $\mathbf{E}$ tal que la representación matricial de T, relativa a esa base, es una matriz diagonal por bloques; donde cada bloque N_j de la diagonal es una matriz nilpotente básica. Además de los bloques nilpotentes N_j hay exactamente:*

$$\left.\begin{array}{l}\bullet\ m_\nu-m_{\nu-1}\geq 1 \text{ bloques de orden } \nu,\\ \bullet\ 2m_{\nu-1}-m_{\nu-2}-m_\nu \text{ bloques de orden } \nu-1,\\ \vdots\\ \bullet\ 2m_2-m_1-m_3 \text{ bloque de orden } 2,\\ \bullet\ 2m_1-m_2 \text{ bloques de orden } 1,\\ \bullet\ m_1, \text{la nulidad de } T, \text{ bloques } N_j \text{ en total.}\end{array}\right\}\tag{D.7}$$

2. *Sea A una matriz cuadrada de orden n que es nilpotente con índice de nilpotencia ν. Sea S_k el espacio nulo de A^k y $m_k=\operatorname{Nul}(A^k)$, la nulidad de A^k, para $k=1,2,\ldots,\nu$. Entonces A es similar a una matriz diagonal por bloques; donde cada bloque N_j de la diagonal es una matriz nilpotente básica y además se cumplen nuevamente las relaciones (D.7) para los bloques N_j (obviamente hay que poner A en lugar de T en la última línea).*

DEMOSTRACIÓN ■ 1. Puesto que el indíce de nilpotencia de T es ν, $S_{\nu-1} \neq \mathbf{E}$ y $S_\nu = \mathbf{E}$; entonces $m_{\nu-1} < m_\nu = n$. Ya que $S_1 \subset S_2 \subset \cdots \subset S_\nu = \mathbf{E}$ (cfr. ejercicio 23, pág. 511), podemos construir, en forma inductiva (cfr. nota al pie de la página 1062), una base $\{\vec{v}_1, \ldots, \vec{v}_n\}$ del espacio $\mathbf{E}$ tal que $\{\vec{v}_1, \ldots, \vec{v}_{m_k}\}$ es una base de S_k para cada k. Generamos ahora, a partir de la base $\{\vec{v}_j\}$ una nueva base en la cual la representación matricial de T tiene la forma deseada mediante el siguiente procedimiento: Sean

$$\begin{aligned}
\vec{u}_{(1,\nu)} &= \vec{v}_{m_{\nu-1}+1} \\
\vec{u}_{(2,\nu)} &= \vec{v}_{m_{\nu-1}+2} \\
&\vdots \\
\vec{u}_{(m_\nu - m_{\nu-1},\nu)} &= \vec{v}_{m_\nu}
\end{aligned}$$

y

$$\begin{aligned}
\vec{u}_{(1,\nu-1)} &= T\vec{u}_{(1,\nu)} \\
\vec{u}_{(2,\nu-1)} &= T\vec{u}_{(2,\nu)} \\
&\vdots \\
\vec{u}_{(m_\nu - m_{\nu-1},\nu-1)} &= T\vec{u}_{(m_\nu - m_{\nu-1},\nu)}.
\end{aligned}$$

Por el lema D.1 $U_1 = \{\vec{v}_1, \ldots, \vec{v}_{m_{\nu-2}}, \vec{u}_{(1,\nu-1)}, \ldots, \vec{u}_{(m_\nu - m_{\nu-1},\nu-1)}\}$ es un subconjunto L.I. de $S_{\nu-1}$. Completemos el conjunto U_1 a una base de $S_{\nu-1}$ adjuntando (si es necesario) elementos denotados por

$$\begin{aligned}
&\vec{u}_{(m_\nu - m_{\nu-1}+1,\nu-1)} \\
&\vec{u}_{(m_\nu - m_{\nu-1}+2,\nu-1)} \\
&\vdots \\
&\vec{u}_{(m_{\nu-1} - m_{\nu-2},\nu-1)}.
\end{aligned}$$

Sean

$$\begin{aligned}
\vec{u}_{(1,\nu-2)} &= T\vec{u}_{(1,\nu-1)} \\
\vec{u}_{(2,\nu-2)} &= T\vec{u}_{(2,\nu-1)} \\
&\vdots \\
\vec{u}_{(m_{\nu-1} - m_{\nu-2},\nu-2)} &= T\vec{u}_{(m_{\nu-1} - m_{\nu-2},\nu-1)}.
\end{aligned}$$

Nuevamente, por el lema D.1, $U_2 = \{\vec{v}_1, \ldots, \vec{v}_{m_{\nu-3}}, \vec{u}_{(1,\nu-2)}, \ldots, \vec{u}_{(m_{\nu-1} - m_{\nu-2},\nu-2)}\}$ es un subconjunto linealmente independiente de $S_{\nu-2}$, que podemos completar (si es necesario) a una base de $S_{\nu-2}$ añadiendo elementos denotados por

$$\begin{aligned}
&\vec{u}_{(m_{\nu-1} - m_{\nu-2}+1,\nu-2)} \\
&\vec{u}_{(m_{\nu-1} - m_{\nu-2}+2,\nu-2)} \\
&\vdots \\
&\vec{u}_{(m_{\nu-2} - m_{\nu-3},\nu-2)}.
\end{aligned}$$

Continuando este proceso obtendremos una base de $\mathbf{E}$ de acuerdo con el siguiente arreglo:

$$\left.\begin{array}{llllllll}
\vec{u}_{(1,\nu)} & \cdots & \vec{u}_{(m_\nu - m_{\nu-1},\nu)} & & & & & \\
\vec{u}_{(1,\nu-1)} & \cdots & \vec{u}_{(m_\nu - m_{\nu-1},\nu-1)} & \vec{u}_{(m_\nu - m_{\nu-1}+1,\nu-1)} & \cdots & \vec{u}_{(m_{\nu-1} - m_{\nu-2},\nu-1)} & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \\
\vec{u}_{(1,2)} & \cdots & \vec{u}_{(m_\nu - m_{\nu-1},2)} & \vec{u}_{(m_\nu - m_{\nu-1}+1,2)} & \cdots & \vec{u}_{(m_{\nu-1} - m_{\nu-2},2)} & \cdots\ \vec{u}_{(m_2 - m_1,2)} & \\
\vec{u}_{(1,1)} & \cdots & \vec{u}_{(m_\nu - m_{\nu-1},1)} & \vec{u}_{(m_\nu - m_{\nu-1}+1,1)} & \cdots & \vec{u}_{(m_{\nu-1} - m_{\nu-2},1)} & \cdots\ \vec{u}_{(m_2 - m_1,1)} & \cdots\ \vec{u}_{(m_1,1)}
\end{array}\right\} \tag{D.8}$$

El último renglón de éste es una base de S_1, los dos últimos renglones forman una base de S_2, etc. Al aplicar T a un elemento de un renglón dado de (D.8), éste se transforma en el elemento del siguiente renglón por debajo de él; es decir,

$$T\vec{u}_{(l,k)} = \begin{cases} \vec{u}_{(l,k-1)} & \text{si } k > 1 \\ \vec{0}_{\mathbf{E}} & \text{si } k = 1 \end{cases} \tag{D.9}$$

Reordenemos la base (D.8) en orden lexicográfico:[9]

$$\mathscr{B} = \left\{\vec{u}_{(1,1)}, \vec{u}_{(1,2)}, \ldots, \vec{u}_{(1,\nu)}, \vec{u}_{(2,1)}, \vec{u}_{(2,2)}, \ldots, \vec{u}_{(m_1,1)}\right\},$$

Consideremos una columna j cualquiera de (D.8), con sus elementos arreglados en orden también lexicográfico y sea $\mathscr{B}_j = \{\vec{u}_{(j,1)}, \ldots, \vec{u}_{(j,\mu)}\}$ la base del subespacio W_j generado por estos vectores; entonces:

(a) W_j es T-invariante. En efecto, si $\vec{w} = \sum_{i=1}^{\mu} \alpha_i \vec{u}_{(j,i)} \in W_j$, por (D.9),

$$\begin{aligned} T\vec{w} &= \sum_{i=1}^{\mu} \alpha_i T\vec{u}_{(j,i)} \\ &= \sum_{i=2}^{\mu} \alpha_i \vec{u}_{(j,i-1)} \in W_j \end{aligned}$$

(b) Sea T_j la restricción del operador T al subespacio W_j; nuevamente por (D.9)

$$[T_j]_{\mathscr{B}_j} = \begin{bmatrix} \vec{0}_{\mathbb{R}^\mu} & \vec{e}_1 & \vec{e}_2 & \cdots & \vec{e}_{\mu-1} \end{bmatrix}$$

donde $\vec{e}_i = (\underbrace{0, 0, \ldots, 0, \overset{i}{1}, 0, \ldots, 0}_{\mu})$. Luego $N_j = [T_j]_{\mathscr{B}_j}$ es una matriz nilpotente básica con índice de nilpotencia μ.

Puesto que $\mathscr{B}$ es la unión disjunta de las bases $\mathscr{B}_j$, $\mathbf{E}$ es la suma directa de los subespacios T-invariantes W_j; esto es, $\mathbf{E} = W_1 \oplus W_2 \oplus \cdots \oplus W_{m_1}$ y, por el teorema D.1, la representación matricial de T es la suma directa de las matrices N_j; esto es

$$[T]_{\mathscr{B}} = N_1 \oplus N_2 \oplus \cdots \oplus N_{m_1} \tag{D.10}$$

Así, la representación matricial del operador T, relativa a la base $\mathscr{B}$, es una matriz diagonal por bloques donde cada bloque es una matriz básica nilpotente. Además, del arreglo (D.8) se desprende que de los bloques nilpotentes N_j hay exactamente:

- $m_\nu - m_{\nu-1}$ de orden ν; y puesto que $m_\nu > m_{\nu-1}$, hay por lo menos uno de orden ν,
- $m_{\nu-1} - m_{\nu-2} - (m_\nu - m_{\nu-1}) = 2m_{\nu-1} - m_{\nu-2} - m_\nu$ de orden $\nu - 1$,
- $2m_2 - m_1 - m_3$ de orden 2 y
- $2m_1 - m_2$ de orden 1.

[9] $(a,b) \preceq (c,d) \Leftrightarrow a < c$ o bien $a = c$ y $b \leq d$ (cfr. el primer inciso de los preliminares al teorema C.5, p. 1046).

Finalmente, como se ve en (D.10)

- el número total de bloques N_j es[10] m_1, la nulidad del operador T.

2. Es consecuencia inmediata del primer inciso al aplicar éste al operador matricial T_A, con $T_A(\vec{u}) = A\vec{u}$ para cada $\vec{u} \in \mathbb{K}^n$. ■

❍ **Nota D.3** Convenimos que en el teorema anterior, y en lo que sigue, $m_0 = 0$ y, en tal caso, un conjunto de la forma $\{\vec{v}_1, \ldots, \vec{v}_{m_0}, \vec{u}_{(j,s)}, \ldots, \vec{u}_{(r,s)}\}$ se interpreta como el conjunto $\{\vec{u}_{(j,s)}, \ldots, \vec{u}_{(r,s)}\}$.

▶ **Ejemplo D.4** Sea $A = \begin{bmatrix} -3 & 1 & 3 & 9 & 4 \\ 1 & -1 & -1 & -5 & -2 \\ -3 & 0 & 3 & 6 & 3 \\ -1 & 0 & 1 & 2 & 1 \\ 1 & 0 & -1 & -2 & -1 \end{bmatrix}$.

1. Comprobar que A es nilpotente y encontrar su índice de nilpotencia.
2. Hallar una matriz N, donde N es una matriz diagonal por bloques y cada bloque es una matriz nilpotente básica, que sea similar a la matriz A.
3. Hallar una matriz invertible Q tal que $Q^{-1}AQ = N$.◀

Solución 1. Se tiene

$$A^2 = \begin{bmatrix} -4 & -4 & 4 & -4 & 0 \\ 2 & 2 & -2 & 2 & 0 \\ -3 & -3 & 3 & -3 & 0 \\ -1 & -1 & 1 & -1 & 0 \\ 1 & 1 & -1 & 1 & 0 \end{bmatrix} \quad \text{y} \quad A^3 = \mathcal{O}.$$

Por tanto, A es nilpotente con índice de nilpotencia $\nu = 3$.

2. Resolviendo el sistema homogéneo $A\vec{x} = \vec{0}$, es fácil demostrar que si S_1 es el espacio nulo de A, entonces

$$\{(1,-1,0,0,1),(2,-3,0,1,0),(1,0,1,0,0)\} \tag{D.11}$$

es una base de S_1. De manera similar, resolviendo el sistema homogéneo $A^2\vec{x} = \vec{0}$, se puede determinar que

$$\{(0,0,0,0,1),(-1,1,0,0,0),(1,0,1,0,0,),(-1,0,0,1,0\} \tag{D.12}$$

es una base de S_2, el espacio nulo de A^2. Utilizando el proceso de compleción de un conjunto L.I. a una base, página 167, podemos completar la base (D.11) a una base de S_2 utilizando los elementos de la base (D.12) para obtener

$$\{(1,-1,0,0,1),(2,-3,0,1,0),(1,0,1,0,0),(0,0,0,0,1)\}$$

[10] El lector puede sumar la cantidad de bloques N_j de cada orden y verificar que efectivamente el número total de ellos es m_1.

Mediante el mismo proceso completamos la base precedente a una base de $S_3 = \mathbb{R}^5$, resultando

$$\{(1,-1,0,0,1),(2,-3,0,1,0),(1,0,1,0,0),(0,0,0,0,1),(1,0,0,0,0)\}$$

la base del espacio nulo de $A^3 = \mathscr{O}$. Entonces, en este caso, $\nu = 3$, $m_1 = 3$, $m_2 = 4$ y $m_3 = 5$. Por tanto,

$$\begin{aligned}\vec{u}_{(1,3)} &= \vec{v}_{m_3-1+1} = \vec{v}_5 = (1,0,0,0,0)\\ \vec{u}_{(1,2)} &= A\vec{u}_{(1,3)} = (-3,1,-3,-1,1)\end{aligned}$$

y

$$\begin{aligned}U_1 &= \{\vec{v}_1,\vec{v}_2,\vec{v}_{m_{3-2}},\vec{u}_{(1,2)}\}\\ &= \{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{u}_{(1,2)}\}\\ &= \{(1,-1,0,0,1),(2,-3,0,1,0),(1,0,1,0,0),(-3,1,-3,-1,1)\}.\end{aligned}$$

Puesto que $\dim(S_2) = 4$, U_1 es una base de S_2 y no hay necesidad de añadir elementos. Por ende,

$$\begin{aligned}\vec{u}_{(1,1)} &= \vec{u}_{(1,\nu-2)}\\ &= A\vec{u}_{(1,\nu-1)}\\ &= A\vec{u}_{(1,2)}\\ &= A(-3,1,-3,-1,1)\\ &= A[\begin{array}{ccccc}-3 & 1 & -3 & -1 & 1\end{array}]^t\\ &= (-4,2,-3,-1,1).\end{aligned}$$

y

$$\begin{aligned}U_2 &= \{\vec{v}_1,\ldots,\vec{v}_{m_{\nu-3}},\vec{u}_{(1,\nu-2)}\}\\ &= \{\vec{u}_{(1,1)}\}\end{aligned}$$

Completemos U_2 a una base de S_1:

$$\{\vec{u}_{(1,1)},\vec{u}_{(m_{\nu-1}-m_{\nu-2}+1,\nu-2)},\vec{u}_{(m_{\nu-1}-m_{\nu-2}+2,\nu-2)}\} = \{\vec{u}_{(1,1)},\vec{u}_{(2,1)},\vec{u}_{(3,1)}\}$$

con

$$\begin{aligned}\vec{u}_{(2,1)} &= \vec{v}_1 = (1,-1,0,0,1),\\ \vec{u}_{(3,1)} &= \vec{v}_2 = (2,-3,0,1,0).\end{aligned}$$

El arreglo, para este caso, que corresponde a (D.8) es

$$\begin{array}{lll}\vec{u}_{(1,3)} & & \\ \vec{u}_{(1,2)} & & \\ \vec{u}_{(1,1)} & \vec{u}_{(2,1)} & \vec{u}_{(3,1)}\end{array}$$

y la base buscada es $\mathscr{B} = \{\vec{u}_{(1,1)},\vec{u}_{(1,2)},\vec{u}_{(1,3)},\vec{u}_{(2,1)},\vec{u}_{(3,1)}\}$. Entonces la matriz en forma canónica nilpotente similar a la matriz A está dada por

$$N = \left[\begin{array}{ccc:c:c} 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ \hdashline 0 & 0 & 0 & 0 & 0\\ \hdashline 0 & 0 & 0 & 0 & 0\end{array}\right]$$

donde hay $m_3 - m_2 = 1$ bloque de orden $\nu = 3$; $2m_2 - m_1 - m_3 = 0$ bloques de orden 2; y $2m_1 - m_2 = 2$ bloques de orden 1.

3. El lector puede verificar, haciendo el producto, que

$$Q^{-1}AQ = N$$

si

$$Q = [\ \vec{u}_{(1,1)} \quad \vec{u}_{(1,2)} \quad \vec{u}_{(1,3)} \quad \vec{u}_{(2,1)} \quad \vec{u}_{(3,1)}\ ] = \begin{bmatrix} -4 & -3 & 1 & 1 & 2 \\ 2 & 1 & 0 & -1 & -3 \\ -3 & -3 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}. \checkmark$$

Un operador lineal en un espacio vectorial de dimensión finita tiene una representación matricial en forma canónica de Jordan si sus polinomios mínimos y característicos se factorizan en factores lineales en el campo $\mathbb{K}$. Por el teorema fundamental del álgebra (teorema B.9, pág. 1025) esto siempre sucede si $\mathbb{K}$ es el campo complejo $\mathbb{C}$. Ya que cualquier campo se puede extender a un campo donde el polinomio característico tenga todas sus raíces (cfr. teorema de Kronecker, página 1026), en un sentido general, todo operador lineal en un espacio de dimensión finita tiene una representación matricial en forma canónica de Jordan. La equivalencia de este hecho a matrices se traduce a que toda matriz, en general, es similar a una forma canónica de Jordan. Establecemos la base de esta afirmación en el siguiente teorema.

Teorema D.6 ***(Formas canónicas de Jordan)***

1. *Sean* $\mathbf{E}$ *un espacio vectorial de dimensión finita y* $T : \mathbf{E} \to \mathbf{E}$ *un operador lineal con sendos polinomios característico y mínimo*

$$p_T(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_r)^{n_r}, \mathfrak{m}_T(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_r)^{m_r};$$

donde los escalares λ_j *son distintos entre sí. Entonces existe una base* $\mathscr{B}$ *de* $\mathbf{E}$ *tal que la representación matricial de* T *relativa a esta base es una forma canónica de Jordan. Específicamente,* $[T]_{\mathscr{B}}$ *es la suma directa de bloques de Jordan* $J_j(\lambda_k)$*,* $k = 1, 2, \ldots, r$*, con las siguientes propiedades:*

$$\left.\begin{array}{l} \text{(a) Existe al menos un bloque } J_j(\lambda_k) \text{ de orden } m_k. \\ \text{(b) La suma de los órdenes de los bloques } J_j(\lambda_k) \text{ es } n_k. \\ \text{(c) El número de bloques } J_j(\lambda_k) \text{ es igual a la multiplicidad geométrica} \\ \quad \text{de } \lambda_k. \\ \text{(d) El número de los } J_j(\lambda_k) \text{ de cada orden posible está únivocamente} \\ \quad \text{determinado por } T. \end{array}\right\} \quad \text{(D.13)}$$

2. *Sea* A *una matriz cuadrada de orden* n *con polinomio característico y polinomio mínimo*

$$p_A(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_r)^{n_r}, \mathfrak{m}_A(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_r)^{m_r},$$

respectivamente; donde los escalares λ_j *son distintos entre sí. Entonces existe una forma canónica de Jordan,* J*, similar a la matriz* A*. Específicamente,* J *es la suma directa de bloques de Jordan* $J_j(\lambda_k)$*,* $k = 1, 2, \ldots, r$*, que cumplen las propiedades (D.13) del inciso precedente (cambiando* T *por* A *en la última condición).*

DEMOSTRACIÓN ■ Por el teorema de descomposición primaria (teorema D.4), T se descompone en las restricciones T_k del operador T al subespacio $S_k = \text{Ker}((T-\lambda_k I)^{m_k})$, donde I es el operador identidad, $I(\vec{u}) = \vec{u}$; esto es, $T = T_1 \oplus T_2 \oplus \cdots \oplus T_r$ y $(\lambda - \lambda_k)^{m_k}$ es el polinomio mínimo de T_k. Entonces $(T_k - \lambda_k I)^{m_k} = \theta$, el operador constante cero, para todo $k = 1, 2, \ldots, r$. Si, para cada k, $N_k = T_k - \lambda_k I$, se tiene que N_k es un operador nilpotente con índice de nilpotencia m_k; del teorema D.5, existe una base $\mathscr{B}_k$ en la cual N_k tiene una representación matricial en forma canónica nilpotente; luego, como

$$T_k = N_k + \lambda_k I,$$

el operador T_k tiene una representación matricial en forma canónica de Jordan, $J_j(\lambda_k)$, relativa a esa misma base. De esta manera, la suma directa J de las matrices $J_j(\lambda_k)$ tiene forma canónica de Jordan y es una representación matricial de T relativa a la reunión de las bases $\mathscr{B}_k$. Finalmente, por el teorema D.5, la propiedad *(a)* se deduce del hecho de que el operador N_k tiene índice de nilpotencia m_k; *(b)* es verdadera porque T y J tienen el mismo polinomio característico; *(c)* se desprende de que la nulidad de $N_k = T_k - \lambda_k I$ es igual a la multiplicidad geométrica del valor propio λ_k; la propiedad *(d)* se desprende del hecho de que los T_k y en consecuencia los N_k, están unívocamente determinados por el operador T. ■

❍ **Nota D.4** Se puede probar que, salvo permutaciones de los bloques en la diagonal, la representación en forma canónica de Jordan de un operador lineal es única. Lo mismo sucede para una matriz en forma canónica de Jordan que sea similar a una matriz dada. Por esta razón nos daremos la libertad de expresar sentencias como “hallar la forma canónica de Jordan de una matriz o de un operador lineal” en lugar de “encontrar una forma canónica de Jordan similar a una matriz” o “hallar una representacion matricial en forma canónica de Jordan para un operador lineal”. Por tanto, si dos matrices son similares, entonces tienen la misma forma canónica de Jordan.

▶ **Ejemplo D.5** Hallar la forma canónica de Jordan de la matriz

$$A = \begin{bmatrix} 6 & 9 & -4 & -4 & 5 \\ -1 & 0 & 1 & 1 & -1 \\ 3 & 6 & -1 & -3 & 2 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \blacktriangleleft$$

Solución Se deja de ejercicio al lector calcular el polinomio característico de la matriz A para obtener $p_A(\lambda) = (\lambda-1)^2(\lambda-2)^3$. Por ende,[11] el polinomio mínimo de A está dado por $\mathfrak{m}_A(\lambda) = (\lambda-1)^2(\lambda-2)^2$. Entonces, en este caso, $\lambda_1 = 1$, $\lambda_2 = 2$, $m_1 = 2$, $m_2 = 2$, $n_1 = 2$ y $n_2 = 3$. Se deja como ejercicio al lector encontrar los vectores propios correspondientes a λ_1 y λ_2 para obtener que

$$\begin{aligned} E_1 &= \text{gn}((5,-1,3,1,0)) \quad \text{y} \\ E_2 &= \text{gn}((1,0,1,0,0),(1,0,0,1,0)) \end{aligned}$$

y, por tanto, la multiplicidad geométrica de λ_1 es $\mu_g(1) = 1$; mientras que la de λ_2 es $\mu_g(2) = 2$. Entonces, utilizando las relaciones (D.13) del teorema D.6:

[11] Cfr. ejercicio 54 del capítulo 5, página 538.

1. El número de bloques $J_j(2)$ es igual a la multiplicidad geométrica de $\lambda_2 = 2$, $\mu_g(2) = 2$; la suma de los órdenes de los bloques $J_j(2)$ es $n_2 = 3$; y debe haber al menos un bloque $J_j(2)$ de orden $m_2 = 2$.
2. El número de bloques $J_j(1)$ es igual a la multiplicidad geométrica de $\lambda_1 = 1$, $\mu_g(1) = 1$; la suma de los órdenes de los bloques $J_j(1)$ es $n_1 = 2$; y debe haber al menos un bloque $J_j(1)$ de orden $m_1 = 2$.

Con esta información la forma canónica de Jordan de la matriz A es:

$$J = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

(La forma canónica de Jordan es única salvo permutaciones de los bloques.) ✔

Una vez que se conoce la forma canónica de Jordan, J, de una matriz dada, A, encontrar una matriz invertible Q tal que $Q^{-1}AQ = J$ —que equivale a encontrar una base en la cual la representación matricial de un operador esté en forma canónica de Jordan— es sencillo cuando el orden de la matriz no es grande. Ilustramos a continuación el procedimiento. Observemos antes que si K es una matriz de tamaño $m \times n$ con columnas $\vec{k}_j$ y $B = [b_{ij}]$ es una matriz de tamaño $n \times p$, con columnas $\vec{b}_j = [\; b_{1j} \quad b_{2j} \quad \cdots \quad b_{jn} \;]^t$ entonces, de (1.2) y (1.5) de los ejemplos 1.14 y 1.15 (páginas 11 y 12), se tiene

$$K[\; \vec{b}_1 \quad \vec{b}_2 \quad \cdots \quad \vec{b}_p \;] = [\; K\vec{b}_1 \quad K\vec{b}_2 \quad \cdots \quad K\vec{b}_p \;]$$
$$\left[\; \sum_{j=1}^{n} b_{j1}\vec{k}_j \quad \sum_{j=1}^{n} b_{j2}\vec{k}_j \quad \cdots \quad \sum_{j=1}^{n} b_{jp}\vec{k}_j \;\right];$$

luego

$$KB = \left[\; \sum_{j=1}^{n} b_{j1}\vec{k}_j \quad \sum_{j=1}^{n} b_{j2}\vec{k}_j \quad \cdots \quad \sum_{j=1}^{n} b_{jp}\vec{k}_j \;\right] \tag{D.14}$$

Ahora supongamos que $J = [b_{ij}]$ es la forma canónica de Jordan y que deseamos encontrar una matriz invertible Q tal que

$$Q^{-1}AQ = J.$$

Sean $\vec{v}_j$ las columnas de la matriz Q, entonces se debe cumplir

$$AQ = QJ.$$

Entonces, por (1.2) y (D.14),

$$[\; A\vec{v}_1 \quad A\vec{v}_2 \quad \cdots \quad A\vec{v}_n \;] = \left[\; \sum_{j=1}^{n} b_{j1}\vec{v}_j \quad \sum_{j=1}^{n} b_{j2}\vec{v}_j \quad \cdots \quad \sum_{j=1}^{n} b_{jn}\vec{v}_j \;\right] \tag{D.15}$$

que da lugar al sistema de ecuaciones

$$A\vec{v}_k = \sum_{j=1}^{n} b_{jk}\vec{v}_j$$

$k = 1, 2, \ldots, n$. Resolviendo este sistema para los $\vec{v}_k$ se puede encontrar la matriz Q. Ilustramos este método en el siguiente ejemplo.

▶ **Ejemplo D.6** Hallar una matriz invertible Q tal que $Q^{-1}AQ = J$ donde A es la matriz del ejemplo precedente y J es su forma canónica de Jordan (la que se calculó en ese ejemplo).◀

Solución Sean $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$, $\vec{v}_4$, y $\vec{v}_5$ las columnas de la matriz Q. Entonces se debe cumplir $AQ = QJ$; es decir,

$$A[\ \vec{v}_1 \quad \vec{v}_2 \quad \vec{v}_3 \quad \vec{v}_4 \quad \vec{v}_5\ ] = [\ \vec{v}_1 \quad \vec{v}_2 \quad \vec{v}_3 \quad \vec{v}_4 \quad \vec{v}_5\ ]J$$

esto es, por (D.15),

$$\begin{aligned}[\ A\vec{v}_1 \quad A\vec{v}_2 \quad A\vec{v}_3 \quad A\vec{v}_4 \quad A\vec{v}_5\ ] &= [\ \vec{v}_1 \quad \vec{v}_2 \quad \vec{v}_3 \quad \vec{v}_4 \quad \vec{v}_5\ ]\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \\ &= [\ \vec{v}_1 \quad \vec{v}_1 + \vec{v}_2 \quad 2\vec{v}_3 \quad \vec{v}_3 + 2\vec{v}_4 \quad 2\vec{v}_5\ ].\end{aligned}$$

Por tanto, se obtiene el sistema

$$\begin{aligned} A\vec{v}_1 &= \vec{v}_1 \\ A\vec{v}_2 &= \vec{v}_1 + \vec{v}_2 \\ A\vec{v}_3 &= 2\vec{v}_3 \\ A\vec{v}_4 &= \vec{v}_3 + 2\vec{v}_4 \\ A\vec{v}_5 &= 2\vec{v}_5 \end{aligned}$$

que equivale a

$$(A - I_5)\vec{v}_1 = \vec{0}_{\mathbb{R}^5} \quad \text{(D.16)}$$

$$(A - I_5)\vec{v}_2 = \vec{v}_1 \quad \text{(D.17)}$$

$$(A - 2I_5)\vec{v}_3 = \vec{0}_{\mathbb{R}^5} \quad \text{(D.18)}$$

$$(A - 2I_5)\vec{v}_4 = \vec{v}_3 \quad \text{(D.19)}$$

$$(A - 2I_5)\vec{v}_5 = \vec{0}_{\mathbb{R}^5} \quad \text{(D.20)}$$

La relación (D.16) entraña que $\vec{v}_1$ debe ser un vector propio correspondiente a $\lambda_1 = 1$; en el ejemplo anterior se encontró $E_1 = \text{gn}((5,-1,3,1,0))$ y, por ende, podemos tomar $\boxed{\vec{v}_1 = (5,-1,3,1,0)}$. Al sustituir $\vec{v}_1$ en (D.17) se obtiene el sistema

$$\left[\begin{array}{rrrrr|r} 5 & 9 & -4 & -4 & 5 & 5 \\ -1 & -1 & 1 & 1 & -1 & -1 \\ 3 & 6 & -2 & -3 & 2 & 3 \\ 0 & 1 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

y al aplicarle el método de Gauss produce el conjunto de soluciones $\{(5 + x_4, -1 - x_4, 4 + 3x_4, x_4, 1) : x_4 \in \mathbb{R}\}$. Elijamos, por simplicidad, $x_4 = 0$ para obtener $\boxed{\vec{v}_2 = (5,-1,4,0,1)}$. La relación (D.18) implica que $\vec{v}_3$ debe ser un vector propio correspondiente a $\lambda_2 = 2$; entonces, por el ejemplo anterior, podemos definir $\boxed{\vec{v}_3 = (1,0,0,1,0)}$. Al sustituir $\vec{v}_3$ en la relación (D.19) resulta el sistema[12]

[12] Se debe elegir el vector propio $\vec{v}_3$ correspondiente al valor propio $\lambda_2 = 2$ de tal manera que el sistema D.19 sea consistente.

$$\left[\begin{array}{rrrrr|r} 4 & 9 & -4 & -4 & 5 & 1 \\ -1 & -2 & 1 & 1 & -1 & 0 \\ 3 & 6 & -3 & -3 & 2 & 0 \\ 0 & 1 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right]$$

que al resolver por el método de Gauss produce el conjunto de soluciones $\{(-2+x_3+x_4,1,x_3,x_4,0): x_3,x_4 \in \mathbb{R}\}$; si elegimos $x_3=0$, $x_4=0$, entonces $\boxed{\vec{v}_4=(-2,1,0,0,0)}$. De la relación (D.20) se desprende que $\vec{v}_5$ debe ser un vector propio correspondiente a $\lambda_2=2$, tomemos $\boxed{\vec{v}_5=(1,0,1,0,0)}$ para que Q sea invertible. De esta manera

$$Q=\left[\begin{array}{rrrrr} 5 & 5 & 1 & -2 & 1 \\ -1 & -1 & 0 & 1 & 0 \\ 3 & 4 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right].$$

El lector puede verificar que, efectivamente, $Q^{-1}AQ=J$. ✔

E | Respuestas a ejercicios seleccionados

Capítulo 1

1.1 Matrices: páginas 55-57

1 (a) $A+B=\begin{bmatrix} 4 & 2 & -4 \\ 0 & 8 & -8 \end{bmatrix}$.

(b) Operación indefinida.

(c) $3A-5B=\begin{bmatrix} -28 & 6 & -4 \\ 16 & 0 & 8 \end{bmatrix}$.

(d) $6D=\begin{bmatrix} 12 \\ 18 \\ -36 \end{bmatrix}$.

3 (a) $A^t=\begin{bmatrix} -1 & 2 \\ 2 & 5 \\ -3 & 4 \end{bmatrix}$.

(b) El producto A^TC no está definido.

(c) $D^tA^t=\begin{bmatrix} 22 & 43 \end{bmatrix}$.

(d) $B^tA=\begin{bmatrix} -9 & 0 & -7 \\ 6 & 15 & -12 \\ -7 & -22 & 19 \end{bmatrix}$.

(e) $D^tD=49$.

(f) $DD^T=\begin{bmatrix} 4 & 6 & -12 \\ 6 & 9 & -18 \\ -12 & -18 & 36 \end{bmatrix}$.

5 (a) La operación no está definida.

(b) La operación no está definida.

(c) $CB=\begin{bmatrix} 4 & -20 & 56 \\ -15 & -5 & -22 \end{bmatrix}$.

7 (a) $A^tB=\begin{bmatrix} -65 & -55 & -17 \\ -8 & 20 & -65 \\ -34 & -10 & -53 \end{bmatrix}$.

(b) $B^tA=\begin{bmatrix} -65 & -8 & -34 \\ -55 & 20 & -10 \\ -17 & -65 & -53 \end{bmatrix}$.

(c) $BB^t=\begin{bmatrix} 90 & 26 \\ 26 & 59 \end{bmatrix}$.

(e) $A^tA=\begin{bmatrix} 73 & 13 & 41 \\ 13 & 106 & 80 \\ 41 & 80 & 74 \end{bmatrix}$.

9 (a) Operación no definida.

(b) $3E-5D=\begin{bmatrix} 22 & 20 & 29 \\ 11 & 19 & -8 \\ 61 & 48 & -4 \end{bmatrix}$.

(c) $-4(D+3E)=\begin{bmatrix} -112 & 16 & -92 \\ 52 & -100 & -64 \\ -52 & 24 & -104 \end{bmatrix}$.

(d) $C-C=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

11 (a) $\frac{1}{4}C^t-\frac{1}{2}A=\begin{bmatrix} -\frac{11}{4} & 1 \\ -\frac{7}{2} & 1 \\ 2 & 5 \end{bmatrix}$.

(b) $B^t-B=\begin{bmatrix} 0 & -9 \\ 9 & 0 \end{bmatrix}$.

(c) $4E^t-3D^t=\begin{bmatrix} 33 & 0 & 52 \\ 12 & 29 & 31 \\ 35 & 4 & 13 \end{bmatrix}$.

(d) $(3E^t-2D^t)^t=\begin{bmatrix} 25 & 8 & 26 \\ -1 & 22 & 4 \\ 37 & 21 & 11 \end{bmatrix}$.

13 (a) $A(BC)=\begin{bmatrix} -140 & 10 & -180 \\ -108 & -206 & -644 \\ -216 & 336 & 480 \end{bmatrix}$.

(b) $CC^t=\begin{bmatrix} 81 & -18 \\ -18 & 56 \end{bmatrix}$.

(c) $(DA)^t=\begin{bmatrix} -31 & -11 & -121 \\ 11 & -29 & -11 \end{bmatrix}$.

(*d*) $(C^t B)A^t \begin{bmatrix} 175 & 387 & -108 \\ -260 & -404 & -96 \\ -360 & -392 & -384 \end{bmatrix}$.

15 (*a*) $(BA^t - 4C)^t = \begin{bmatrix} -51 & 4 \\ -91 & -60 \\ -68 & 40 \end{bmatrix}$.

(*b*) $B^t(CC^t - A^t A) = \begin{bmatrix} 104 & 556 \\ -107 & 276 \end{bmatrix}$.

(*c*) $D^t E^t - (ED)^t = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

17 (*a*) $A^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

(*b*) $A^7 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2187 & 0 \\ 0 & 0 & 128 \end{bmatrix}$.

19 (*b*) $B = \begin{bmatrix} 2b & b \\ 2d & d \end{bmatrix}, b, d \in \mathbb{R}$.

21 $\alpha = 0$ y $\beta = -2$.

23 $\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}^n = \begin{bmatrix} \cos n\theta & -\operatorname{sen} n\theta \\ \operatorname{sen} n\theta & \cos n\theta \end{bmatrix}$.

25 $\begin{bmatrix} a & b \\ -\frac{a^2}{b} & -a \end{bmatrix}, a, b \in \mathbb{R}, b \neq 0; \begin{bmatrix} a & -\frac{a^2}{c} \\ c & -a \end{bmatrix}, a, c \in \mathbb{R}, c \neq 0; \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

27 $\begin{bmatrix} a & b \\ \frac{1-a^2}{b} & -a \end{bmatrix}, a, b \in \mathbb{R}, b \neq 0; \begin{bmatrix} a & \frac{1-a^2}{c} \\ c & -a \end{bmatrix}, a, c \in \mathbb{R}, c \neq 0; \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$.

29 $\begin{bmatrix} 1 & -2 & \boxed{1} & 3 \\ \boxed{-2} & 3 & \boxed{-2} & 6 \\ 1 & -2 & 4 & \boxed{1} \\ \boxed{3} & \boxed{6} & 1 & 5 \end{bmatrix}$.

32 Verdadera.

34 Verdadera.

36 Falsa.

38 Falsa.

1.2 Sistemas lineales: páginas 57-62

40 (*a*) Lineal.

(*b*) No lineal.

(*c*) No lineal.

(*d*) Lineal.

42 (*a*) Forma escalonada. Pivotes: 2 (primera fila, primera columna); -1 (segunda fila); 2 (tercera fila).

(*b*) No está en forma escalonada. El primer elemento no nulo de la tercera fila no está a la derecha del primer elemento no nulo de la segunda fila.

(*c*) Forma escalonada. Pivotes: -1 (primera fila); 1 (segunda fila).

(*d*) No está en forma escalonada. La primera componente no nula de la tercera fila no está a la derecha de la primera componente no nula de la segunda fila.

45 (*a*) $\left[\begin{array}{ccccc|c} 1 & -1 & 1 & -1 & -1 & -6 \\ 0 & 1 & -1 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 2 & -8 \end{array}\right]$.

(*b*) Está en forma escalonada.

(*c*) Variables ligadas: x_1, x_2, x_5; variables libres: x_3, x_4.

46 (*a*) $\left[\begin{array}{ccccc|c} 1 & -4 & 5 & -1 & 1 & -7 \\ 0 & 0 & 0 & 1 & -5 & 12 \\ 0 & 0 & 0 & 0 & 3 & -4 \end{array}\right]$.

(*b*) Está en forma escalonada.

(*c*) Variables ligadas: x_1, x_4, x_5; variables libres: x_2, x_3.

49 Inconsistente.

50 Tiene una infinidad de soluciones. Variables ligadas: x_1, x_3, x_4; variables libres: x_2.

51 (*c*) $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -9+r \\ r \\ \frac{9}{2} \\ 2 \end{bmatrix}, r \in \mathbb{R}$.

53 $H = \begin{bmatrix} 1 & -1 & 2 & 1 & 5 \\ 0 & 1 & -3 & 0 & -7 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & -6 & -7 \end{bmatrix}$ (la respuesta no es única).

55 $H = \begin{bmatrix} 0 & 1 & -3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

57 $H = \begin{bmatrix} -1 & 2 & 3 & -1 & 1 & 0 \\ 0 & -1 & -10 & 4 & -6 & 2 \\ 0 & 0 & -39 & 18 & -26 & 18 \\ 0 & 0 & 0 & -147 & 130 & -615 \end{bmatrix}$.

59 $H = \begin{bmatrix} 1 & 2 & -5 & 0 & -2 & -4 \\ 0 & 1 & -18 & -4 & -11 & -32 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 & 8 \\ 0 & 0 & 0 & 0 & 1 & 57 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

61 $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -\frac{5}{8} - 7r - 3s \\ \frac{9}{8} - 5r + 11s \\ 8r \\ 4s \end{bmatrix}; r, s \in \mathbb{R}$.

63 $\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6 \end{bmatrix} = \begin{bmatrix} \frac{61}{27}-17s\\ \frac{656}{27}-274s-135r\\ \frac{47}{3}-144s-81r\\ 27r\\ -\frac{77}{27}+52s\\ 27s \end{bmatrix}$; $r,s\in\mathbb{R}$.

65 $\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{bmatrix} = \begin{bmatrix} 3-r\\ -3+2r\\ r\\ 1 \end{bmatrix}$; $r\in\mathbb{R}$.

67 Inconsistente.

69 $\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{bmatrix} = \begin{bmatrix} 1\\ 2\\ 3\\ 4 \end{bmatrix}$.

71 $\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{bmatrix} = \begin{bmatrix} 1\\ -1\\ 5\\ 2 \end{bmatrix}$.

73 $\begin{bmatrix} x_1\\ x_2\\ x_3 \end{bmatrix} = \begin{bmatrix} 1\\ 0\\ 2 \end{bmatrix}, \begin{bmatrix} -1\\ 2\\ 1 \end{bmatrix}, \begin{bmatrix} -3\\ -1\\ 2 \end{bmatrix}$.

75 $\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{bmatrix} = \begin{bmatrix} 3-2r\\ -3+2r\\ 0\\ r \end{bmatrix}, \begin{bmatrix} 3-2r\\ -1+2r\\ -1\\ r \end{bmatrix}, \begin{bmatrix} -3-2r\\ 3+2r\\ 0\\ r \end{bmatrix}$; $r\in\mathbb{R}$.

77 $H = \begin{bmatrix} 1 & 0 & 0 & -\frac{5}{6} & 0\\ 0 & 1 & 0 & \frac{85}{12} & \frac{5}{2}\\ 0 & 0 & 1 & \frac{175}{24} & \frac{13}{4} \end{bmatrix}$.

79 $H = \begin{bmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 \end{bmatrix}$.

81 $H = \begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1 \end{bmatrix}$.

83 $H = \begin{bmatrix} 1 & 3 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}$.

85 $H = \begin{bmatrix} 1 & 2 & 0 & 0 & 3 & -6\\ 0 & 0 & 1 & 0 & 1 & -2\\ 0 & 0 & 0 & 1 & -1 & 3 \end{bmatrix}$.

87 $\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{bmatrix} = \begin{bmatrix} 29-2s+14r\\ s\\ 12+5r\\ r \end{bmatrix}$; $r,s\in\mathbb{R}$.

89 $\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{bmatrix} = \begin{bmatrix} -1\\ 1\\ 0\\ 1 \end{bmatrix}$.

91 $\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5 \end{bmatrix} = \begin{bmatrix} 5+r_1-r_2+2r_3\\ r_1\\ -1\\ r_2\\ r_3 \end{bmatrix}$; $r_1,r_2,r_3\in\mathbb{R}$.

93 $\vec{x}=(0,0,0)$.

95 $\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{bmatrix} = \begin{bmatrix} -11r\\ r\\ -5r\\ 8r \end{bmatrix}$; $r\in\mathbb{R}$.

97 $\vec{x}=\vec{p}+\vec{h}=\begin{bmatrix} -5/8\\ 7/4\\ 0\\ 0 \end{bmatrix}+\begin{bmatrix} -s+r\\ 2s+2r\\ 4s\\ 8r \end{bmatrix}$; $r,s\in\mathbb{R}$.

99 $\vec{x}=\vec{p}+\vec{h}=\begin{bmatrix} 5/3\\ 2/3\\ 0\\ 0 \end{bmatrix}+\begin{bmatrix} -5s-r\\ s+2r\\ 3s\\ 3r \end{bmatrix}$; $r,s\in\mathbb{R}$.

101 $\vec{x}=\vec{p}+\vec{h}=\begin{bmatrix} 10/11\\ 4/11\\ 0\\ 0 \end{bmatrix}+\begin{bmatrix} -4s-13r\\ 5s-3r\\ 11s\\ 11r \end{bmatrix}$; $r,s\in\mathbb{R}$.

102 $\alpha\neq\pm4$.

103 $\alpha=-4$.

104 $\alpha=4$.

108 $3\alpha-\beta+\gamma=0$.

112 $\vec{z}=(1,-i,i)$.

114 $\begin{bmatrix} z_1\\ z_2\\ z_3\\ z_4 \end{bmatrix} = \begin{bmatrix} \frac{2}{5}+\frac{6}{5}i\\ -\frac{3}{10}+\frac{1}{10}i\\ -\frac{3}{10}+\frac{1}{10}i\\ \frac{3}{5}-\frac{1}{5}i \end{bmatrix}$.

Capítulo 2

2.1 Matrices invertibles y sus inversas:
páginas 102-106

1 $\begin{bmatrix} 1 & 1\\ 0 & -1 \end{bmatrix}$.

3 $\begin{bmatrix} 1 & 0 & 1\\ 0 & \frac{1}{2} & \frac{1}{2}\\ 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$.

5 Matriz singular (no invertible).

7 $\begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -1 & 1 & 1 \end{bmatrix}.$

8 1. $\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$

3. $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$

9 $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$

11 $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}.$

13 $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}.$

15 $\begin{bmatrix} -4 & -1 \\ -\frac{3}{2} & -\frac{1}{2} \end{bmatrix}.$

17 Matriz singular (no invertible).

19 $\begin{bmatrix} -4 & 1 & 2 \\ 1 & 1 & 0 \\ 3 & 0 & -1 \end{bmatrix}.$

21 $\begin{bmatrix} -1 & 2 & 7 \\ 1 & -1 & -5 \\ -2 & 3 & 13 \end{bmatrix}.$

23 $\begin{bmatrix} -\frac{3}{2} & 1 & \frac{1}{2} \\ -\frac{5}{2} & 2 & -\frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}.$

25 $\begin{bmatrix} -\frac{35}{3} & \frac{4}{3} & -2 \\ -3 & 0 & -1 \\ \frac{5}{3} & -\frac{1}{3} & 0 \end{bmatrix}.$

27 $\begin{bmatrix} -12 & 2 & -1 & 5 \\ -2 & 1 & 0 & 1 \\ -3 & 0 & 1 & 0 \\ 4 & 0 & 0 & -1 \end{bmatrix}.$

29 $\begin{bmatrix} 13 & 2 & -3 & 1 \\ 3 & 1 & 0 & 0 \\ 2 & 0 & -1 & 0 \\ 10 & 0 & -4 & 1 \end{bmatrix}.$

31 $\begin{bmatrix} 13 & -7 & -2 & -17 \\ 7 & -4 & -1 & -9 \\ -31 & 18 & 5 & 40 \\ -19 & 11 & 3 & 24 \end{bmatrix}.$

33 Matriz singular (no invertible).

35 $\begin{bmatrix} -34 & 1 & -2 & 3 & 4 \\ -2 & 1 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 \\ 3 & 0 & 0 & -1 & 0 \\ 5 & 0 & 0 & 0 & -1 \end{bmatrix}.$

37 $\begin{bmatrix} 0 & \frac{1}{2} & 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 3 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 9 & 0 & -3 & 0 & 1 & 0 \end{bmatrix}.$

39 $B = \begin{bmatrix} 6 & -6 & 5 \\ -15 & 6 & 7 \\ 6 & 6 & -12 \end{bmatrix}.$

41 $X = \begin{bmatrix} -3 & -5 \\ 2 & 0 \end{bmatrix}.$

43 $\alpha \in \mathbb{R} - \{-2\}.$

46 $(A-I)^{-1} = -(A+I).$

47 Verdadera.

49 Verdadera.

51 Falsa.

53 Verdadera.

55 Falsa.

57 Falsa.

61 $\begin{bmatrix} -4+i & 1-i & -i & -2i \\ -2i & i & 0 & 0 \\ -1-i & 0 & -i & 0 \\ 1 & 0 & 0 & i \end{bmatrix}.$

2.2 Determinantes, adjunta y regla de Cramer:
páginas 106-110

63 27.

65 51.

67 (a) 399

(b) −174

(c) −123.

70 28.

72 33.

74 76.

76 −459.

81 240.

83 −72.

85 6.

87 273.

89 $-1\,060$.

91 16.

93 $-1/4$.

95 -16.

102 (a) $\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix}$

(b) $\begin{bmatrix} 2 & 4 \\ -3 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{8} & -\frac{1}{4} \\ \frac{3}{16} & \frac{1}{8} \end{bmatrix}$

(c) $\begin{bmatrix} 3 & 6 \\ 1 & 2 \end{bmatrix}^{-1}$ no existe.

104 $A^{-1} = \begin{bmatrix} 3 & 1 & -2 \\ 7 & 3 & -5 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$.

106 $A^{-1} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & \frac{1}{2} & -2 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$.

111 $\lambda = -1$ y $\lambda = -4$.

113 $\lambda = \frac{5}{2} \pm \frac{1}{2}\sqrt{41}$.

115 $x_1 = \frac{19}{5}, x_2 = \frac{18}{5}$.

117 $x_3 = -3, x_2 = -2, x_1 = -1$.

119 $x_4 = 2$.

121 $1 + 11i$.

123 $-2 + 7i$.

Capítulo 3

3.1 Geometría de los espacios $\mathbb{R}^n$: páginas 207-211

1 $-\vec{u} = (1, -2, -4)$.

3 $\vec{u} + \vec{v} = (-2, 2, 6)$.

5 $\vec{v} - 5\vec{w} = (14, -20, -23)$.

7 $\vec{w} - \vec{u} = (-2, 2, 1)$.

10 $\vec{u} \cdot (3\vec{v} + \vec{w}) = 58$.

12 $\arccos\left(\frac{3}{35}\sqrt{5}\sqrt{21}\right) \approx 0.498\,48$.

14 $x = 14$.

16 $(2, -1, 1)$ (y cualquier múltiplo escalar no nulo de este vector).

18 $\vec{u} \times \vec{v} = (4, -2, 2)$; $\vec{v} \times \vec{u} = (-4, 2, -2)$.

20 $(\vec{u} \times \vec{w}) \times \vec{v} = (-14, 10, -7)$.

26 26 unidades de área.

29 120 unidades de volumen.

32 17 unidades de volumen.

34 18 unidades de área.

36 $\sqrt{390}$ unidades de área.

38 $\sqrt{110}$ unidades de área.

40 $-\vec{w} = (-2, 1, 3, -1, 2)$.

42 $\vec{u} + \vec{w} = (3, -3, -7, 4, 3)$.

44 $\vec{v} - 3\vec{w} = (-8, 6, 8, -1, 5)$.

46 $\vec{w} - \vec{u} = (1, 1, 1, -2, -7)$.

48 $-\vec{v}$ (la respuesta no es única).

50 $\vec{u} \cdot \vec{v} = -3$.

52 $(\vec{u} + \vec{v}) \cdot \vec{w} = 9$.

54 $\frac{\pi}{2}$.

56 $y = -15$.

58 $(6, 4, 3, 3, 1)$ (la respuesta no es única).

60 $\sqrt{29}$.

62 $\sqrt{133}$.

69 $-3x + 2y + 4z = 27$.

73 $7x + 29y + 10z = 45$.

3.2 Espacios vectoriales: páginas 211-219

77 (a) No es cerrada.
(b) No es cerrada.

79 (a) Es cerrada.
(b) Es cerrada.

81 (a) Es cerrada.
(b) Es cerrada.

83 (a) Es cerrada.
(b) Es cerrada.

85 (a) No es cerrada.
(b) No es cerrada.

87 (a) No es cerrada.
(b) No es cerrada.

89 (a) Es cerrada.
(b) Es cerrada.

90 No es un espacio vectorial. No se cumplen las propiedades: 7 y 10.

92 No es un espacio vectorial. No se cumple la propiedad 10.

94 Sí es un espacio vectorial.

96 No es un espacio vectorial. No se cumple la propiedad 9.

103 $S < \mathbb{R}^2$.

105 S no es un subespacio de $\mathbb{R}^2$.

107 $S < \mathbb{R}^4$.

109 S no es un subespacio de $\mathbb{R}^2$.

111 S no es un subespacio de $\mathfrak{M}_n(\mathbb{R})$.

113 $S < \mathbb{R}^\infty$.

115 $S < \mathbb{R}^\infty$.

130 *(i)* $S_1 \cap S_2 = \{\vec{0}_{\mathbb{R}^3}\}$.

138 El plano $x+y+z=0$.

144 $\vec{v} = -\frac{12}{5}\vec{u}_1 - \frac{1}{5}\vec{u}$.

146 $\vec{v} = \frac{3}{4}\vec{u}_1 - \frac{7}{4}\vec{u}_2 + \frac{7}{4}\vec{u}_3$.

148 $\vec{v} = 2\vec{u}_1 - 4\vec{u}_2 + \vec{u}_4$.

150 $\vec{v} = \vec{u}_1 - \vec{u}_2$.

152 $\vec{v} = 2\vec{u}_1 - \vec{u}_2$.

154 $\vec{v} = -\frac{1}{2}\vec{u}_1 - 10\vec{u}_2 + 3\vec{u}_3$.

156 $\vec{v} = \sum_{i=1}^{n} \vec{u}_i$.

158 $\vec{v} = \frac{1}{2}\vec{u}_1 + \frac{1}{2}\vec{u}_2$.

160 $\vec{v} - 3\vec{u}_1 + 2\vec{u}_2 - \vec{u}_3 \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$.

162 Sí, $\text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3) = \mathbb{R}^3$.

165 No, $\text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3) \neq \mathbb{R}^3$.

166 $\vec{v} = \vec{u}_1 - \vec{u}_2 - 2\vec{u}_3 + \vec{u}_4 \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4)$.

167 $\vec{v} = \vec{u}_1 - \vec{u}_2 - \vec{u}_3 + \vec{u}_4 \in \text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4)$.

168 $\text{gn}(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4) = \mathfrak{M}_2$.

173 Sí generan a $\mathbb{R}^2$.

175 Sí generan a $\mathbb{R}^3$.

177 Sí generan a $\mathbb{R}^3$

179 Sí generan a $\mathbb{R}^4$

181 Sí, $\text{gn}((-1,2)\ ,(-1,1)) = \mathbb{R}^2$.

182 No generan a $\mathbb{R}^3$.

184 No generan a $\mathbb{R}^3$.

186 Sí, $\text{gn}((1,-2,1,2), (-1,0,2,3), (-2,1,3,2), (-1,1,1,0)) = \mathbb{R}^4$.

188 No generan a $\mathbb{R}^4$.

190 $\mathbf{P}_2$

192 El subespacio de matrices triangulares superiormente en $\mathfrak{M}_2(\mathbb{R})$.

194 $\mathbf{P}$.

3.3 Dependencia e independencia lineales: páginas 219-222

211 L.D.

213 L.D.

215 L.D.

217 L.I.

219 L.D.

221 L.D.

223 L.I.

225 L.I.

227 L.I.

229 L.I.

231 L.I.

233 L.D.

235 L.D.

237 L.D.

239 L.I.

241 L.D.

243 L.I.

245 L.I.

247 L.D.

249 L.D.

251 L.I.

253 L.I.

255 L.I.

257 L.I.

261 $\alpha \in \mathbb{R} - \{10\}$.

3.4 Bases y dimensión: páginas 222-231

269 L.D.

271 L.I.

274 Sí es base.

276 No es base.

278 Sí es base.

280 No es base.

282 No es base.

284 No es base.

286 Sí es base.

288 Sí es base.

290 Sí es base.

292 No es base.

294 Sí es base.

296 Sí es base.

299 $S < \mathbb{R}^3$, $\{(0,1,0),(0,0,1)\}$, $\dim(S) = 2$.

301 $S < \mathbb{R}^3$, $\{(-1,0,1),(-1,1,0)\}$, $\dim(S) = 2$.

303 $S < \mathbb{R}^3$, $\{(1,1,0),(0,0,1)\}$, $\dim(S) = 2$.

305 S no es un subespacio de $\mathbb{R}^3$.

307 $S < \mathbf{P}_n$, $\{1\}$, $\dim(S) = 1$.

309 $S < \mathbf{P}_n$, $\{-x+x^2, -x+x^3, \ldots, -x+x^n\}$, $\dim(S) = n-1$.

312 $S < \mathbf{P}_n$. Si $n = 2k$ o $n = 2k+1$, $k = 0,1,\ldots$, una base es $\{1, x^2, x^4, \ldots, x^{2k}\}$ y $\dim(S) = k+1$.

316 Si $\alpha = 0$ o $\beta = 0$, $\dim(S) = 2$. Si $\alpha \neq 0$ y $\beta \neq 0$, $\dim(S) = 3$.

318 Si $\alpha = 0$, $\dim(S) = 2$; si $\alpha \neq 0$, $\dim(S) = 3$.

320 $\dim(S) = 2$.

322 $\dim(S) = 2$.

324 $\dim(S) = 3$.

327 $\{(3,6),(-1,3)\}$, $\dim(S) = 2$.

329 $\{(1,-1,2),(3,-1,1)\}, \dim(S) = 2$.

331 $\{(1,1,1),(-2,1,3),(-1,2,1)\}$, $\dim(S) = 3$.

333 $\{(2,1,3,2),(-1,1,-1,1)\}$, $\dim(S) = 2$.

335 $\left\{\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 3 & -1 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}\right\}$, $\dim(S) = 3$.

337 $\left\{\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}$, $\dim(S) = 4$.

339 $\{2, 3+2x, x^2+1\}$, $\dim(S) = 3$.

341 $\{2-x, 2+x, x-x^2, x^3+1, x^4-x\}$, $\dim(S) = 5$.

343 $\{\text{sen } 4x, \text{sen } 2x\}$, $\dim(S) = 2$.

345 $\{1, e^x, e^{-x}\}$, $\dim(S) = 3$.

348 $\{1+4x-2x^2+x^3, 1-9x+3x^2-2x^3, 5+7x-5x^2+3x^3\}$, $\dim(S) = 3$.

350 $\{2+x+x^2-3x^3, 5-2x+7x^2-9x^3, -1+2x^2\}$, $\dim(S) = 3$.

352 *(a)* Para $n = 2$: $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$, $\dim(S) = 3$.

353 *(b)* $\dim(S) = \frac{n(n+1)}{2}$.

355 $\{M_{kl} = [m_{ij}] \in \mathfrak{M}_n(\mathbb{R}) : k,l \in \{1,2,\ldots,n\}, k \neq l\}$, donde $m_{ij} = \begin{cases} 1, & \text{si } i = k \text{ y } j = l \\ 0, & \text{en otro caso} \end{cases}$ $\dim(S) = n^2 - n$.

356 Sea $k = \dim(S)$ y $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ una base de S, luego, si $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-k}\}$ son vectores en $\mathbf{E} - S$ tales que $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-k}\}$ es una base de $\mathbf{E}$, entonces $\{[\vec{v}_1], [\vec{v}_2], \ldots, [\vec{v}_{n-k}]\}$ es una base de $\mathbf{E}/S$ y, por tanto, $\dim(\mathbf{E}/S) = n-k$.

359 $\{(-1,2,3),(2,-1,2),(1,0,0)\}$.

361 $\{(2,-1,2,2),(-1,2,3,-2),(1,0,0,0),(0,1,0,0)\}$.

363 $\{(-1,0,2,3,2),(-1,1,2,3,-2),(-1,1,0,1,1), (1,0,0,0,0),(0,1,0,0,0)\}$.

365 $\left\{\begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right\}$.

367 $\left\{\begin{bmatrix} -1 & 2 & 1 \\ 2 & -1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} -2 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right\}$.

369 $\mathscr{B} = \{2, x-2, (x-2)^2, x^3, x^4\}$.

376 Para S_1: $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}$, $\dim(S_1) = 3$. Para S_2: $\left\{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$,
$\dim(S_2) = 2$. Para $S_1 + S_2$: $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$, $\dim(S_1 + S_2) = 4$.
Para $S_1 \cap S_2$: $\left\{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right\}$, $\dim(S_1 \cap S_2) = 1$.

378 $\{(-15,-16,4,0,1),(9,8,-3,1,0)\}$, $\dim(S) = 2$.

380 $\dim(S) = 0$.

382 $\{(-140,-139,29,-31,0,268),(16,-9,-55,38,67,0)\}$, $\dim(S) = 2$.

384 Espacio fila: $\{(1,-2,3),(0,1,4),(0,0,5)\}$, $\dim(E_f) = 3$.
Espacio columna: $\{(1,-1,4,0),(-2,3,-1,-1),(3,1,2,1)\}$, $\dim(E_c) = 3$. $\dim(\text{Nul}(A)) = 0$. $\text{Rang}(A) = 3$.

386 Espacio fila: $\{(1,-1,2,3),(0,-1,3,5),(0,0,4,4)\}$, $\dim(E_f) = 3$.
Espacio columna: $\{(1,-1,-1,1),(-1,0,1,-1),(2,1,2,6)\}$, $\dim(E_c) = 3$.
Espacio nulo: $\{(1,2,-1,1)\}$, $\dim(\text{Nul}(A)) = 1$. $\text{Rang}(A) = 3$.

389 $$\begin{aligned} -13x - 11y - z + 8w &= 0 \\ 19x + 5y - 9z + 8t &= 0 \end{aligned}$$

393 Para $S_1 + S_2$: $\{(-1,2,-1,2,2),(2,-1,2,3,1),(3,-1,2,1,0),(-2,3,1,2,1),(-1,1,1,2,1)\}$ y $\dim(S_1 + S_2) = 5$.
Para $S_1 \cap S_2$: $\{(-10,-55,53,139,42)\}$, $\dim(S_1 \cap S_2) = 1$.

396 Para $S_1 + S_2$: $\{1 - x + x^2 + 2x^3, 2 - 3x + 2x^2 - x^3, -1 + 2x + 3x^2 - 4x^3, -2 + 3x - 4x^2 + x^3\}$ y $\dim(S_1 + S_2) = 4$. Para $S_1 \cap S_2$: $\{308 - 475x + 320x^3, 244 - 315x + 320x^2\}$ y $\dim(S_1 \cap S_2) = 2$.

3.5 Espacios vectoriales complejos: páginas 232-233

404 $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & i \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ i & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & i \end{bmatrix}\right\}$, dimensión 8.

406 $\{(1,-2i,2i),(1,-1,2),(1,0,0)\}$

408 $\left\{\begin{bmatrix} 1 & -i & 1 \\ 2 & -1 & 1 \end{bmatrix}, \begin{bmatrix} -2 & 2 & -i \\ 1 & 0 & i \end{bmatrix}\right\}$, $\dim(S) = 2$.

Capítulo 4

4.1 Espacios con producto interior: páginas 383-400

1 No es un producto interior. No se cumplen las propiedades 4 y 5.

3 No es producto interior. No se cumplen las cinco propiedades.

21 *(a)* -212. *(b)* $2\sqrt{41}$. *(c)* $2\sqrt{5}$.

23 *(a)* -3. *(b)* $\sqrt{11}$. *(c)* $\sqrt{39}$.

25 $\begin{bmatrix} 5 & -11 \\ -11 & 25 \end{bmatrix}$.

27 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.

29 $\begin{bmatrix} 2 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{bmatrix}$.

32 No es un producto interior; sólo cumple la propiedad 2.

34 No es un producto interior; no cumple las propiedades 2 y 3.

36 Sí es un producto interior.

38 1.

40 33.

42 1.

44 $\left(-\frac{4}{3}, -\frac{4}{3}, \frac{4}{3}\right)$.

46 $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

48 $\begin{bmatrix} -\frac{3}{2} \\ 0 \\ \frac{3}{2} \end{bmatrix}$

50 0.

52 -11.

54 0.

56 $(0,0,0,0)$.

58 $\begin{bmatrix} \frac{11}{17} & -\frac{1}{17} & \frac{4}{17} & \frac{7}{17} \\ -\frac{1}{17} & \frac{14}{17} & -\frac{5}{17} & \frac{4}{17} \\ \frac{4}{17} & -\frac{5}{17} & \frac{3}{17} & \frac{1}{17} \\ \frac{7}{17} & \frac{4}{17} & \frac{1}{17} & \frac{6}{17} \end{bmatrix}$.

60 0.

62 $2e^{-1}$.

64 $\left(\frac{1}{2}e^2 - \frac{1}{2}e^{-2}\right)^{1/2}$.

66 $\frac{1}{15}\sqrt{16}\sqrt{15}$.

68 $\frac{1}{2}\pi$.

70 $p(x) = \sqrt{6}e^{-1}x, -1 \leq x \leq 1$.

72 $p(x) = \frac{1}{2}e - \frac{1}{2}e^{-1} + 3e^{-1}x, x \in [-1,1]$.

74 $\sqrt{6}$.

76 $\sqrt{19}$.

78 $0.869122 \approx 49.8°$.

80 $p = -\frac{2}{19}\begin{bmatrix} -1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

82 $p = \begin{bmatrix} \frac{2}{3} & -1 & \frac{1}{3} \\ 1 & -\frac{5}{3} & 2 \end{bmatrix}$.

84 $\langle A,B \rangle = \frac{1}{2}$.

86 $\langle A,B \rangle = \frac{\sqrt{2}}{2} - 1$.

88 $\|A\| = \left(\frac{1}{\sqrt{5}-1}\right)^{1/2}$.

90 Sugerencia: Utilizar el hecho de que $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{6}\pi^2$. $A = \left(\frac{1}{4n+2}\right)$, $B = \left(\frac{1}{4n-2}\right)$; $\langle A,B \rangle = \frac{1}{8}$;

$\|A\| = \sqrt{\frac{1}{32}\pi^2 - \frac{1}{4}}$, $\|B\| = \frac{\sqrt{2}}{8}\pi$; $\theta = \arccos\left(\dfrac{4}{\pi\sqrt{\pi^2-8}}\right) \approx 0.37315 \approx 21.40°$.

92 $p = -\frac{15}{13}B$

102 $S^{\perp} = \text{gn}(4,3,-2)$.

104 $S^{\perp} = \text{gn}(5,5,7,-2)$.

106 $S^{\perp} = \text{gn}\left(\left\{\begin{bmatrix} 7 & 6 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ -1 & 0 \end{bmatrix}\right\}\right)$.

108 $S^{\perp} = \text{gn}(2x-1)$.

121 $\left\{\left(-\sqrt{2}/2, \sqrt{2}/2,\right), \left(\sqrt{2}/2, \sqrt{2}/2\right)\right\}$.

123 $\{(0,1),(1,0)\}$.

125 $\left\{\left(\frac{3}{11}\sqrt{11}, \frac{1}{11}\sqrt{11}, \frac{1}{11}\sqrt{11}\right), \left(-\frac{1}{11}\sqrt{22}, \frac{3}{22}\sqrt{22}, \frac{3}{22}\sqrt{22}\right)\right\}$.

127 $\left\{\left(-\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}, 0\right), \left(\frac{1}{6}\sqrt{2}, \frac{1}{6}\sqrt{2}, \frac{2}{3}\sqrt{2}\right), \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right)\right\}$.

129 $\left\{\left(-\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}, 0\right), \left(\frac{1}{6}\sqrt{6}, \frac{1}{6}\sqrt{6}, \frac{1}{3}\sqrt{6}\right), \left(\frac{1}{3}\sqrt{3}, \frac{1}{3}\sqrt{3}, -\frac{1}{3}\sqrt{3}\right)\right\}$.

131 $\left\{\left(\frac{1}{2}\sqrt{2}, 0, \frac{1}{2}\sqrt{2}, 0\right), (0,1,0,0), \left(\frac{1}{6}\sqrt{6}, 0, -\frac{1}{6}\sqrt{6}, \frac{1}{3}\sqrt{6}\right)\right\}$.

133 $\left\{\sqrt{\frac{2}{\pi}}\,\text{sen}\, x, \sqrt{\frac{2}{\pi}}\cos x\right\}$.

135 $\left\{1, \sqrt{3}\,(2x-1), \sqrt{5}\,(6x^2-6x+1)\right\}$.

137 $\left\{\frac{\sqrt{3}}{3}\begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}, \frac{\sqrt{51}}{51}\begin{bmatrix} 4 & -1 \\ 5 & 3 \end{bmatrix}\right\}$.

139 $\left\{\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \frac{\sqrt{7}}{14}\begin{bmatrix} 1 & -3 & 1 \\ -4 & 1 & 0 \end{bmatrix}, \frac{\sqrt{119}}{119}\begin{bmatrix} 6 & -4 & -1 \\ 4 & -1 & 7 \end{bmatrix}\right\}$.

141 $\left\{\left(\frac{9}{1106}\sqrt{1106}, \frac{2}{79}\sqrt{1106}, \frac{2}{553}\sqrt{1106}, -\frac{15}{1106}\sqrt{1106}\right)\right\}$.

143 $\left\{\left(\frac{4}{37}\sqrt{74}, -\frac{1}{74}\sqrt{74}, \frac{3}{74}\sqrt{74}\right)\right\}$.

144 $\left\{\sqrt{3}\left(\frac{1}{2^n}\right), \sqrt{200}\left(\left(\frac{1}{3^n}\right) - \frac{3}{5}\left(\frac{1}{2^n}\right)\right)\right\}$.

146 0.

148 0.

150 $\frac{1}{3}\sqrt{3}\sqrt{8}$.

152 $\frac{1}{2}\pi$.

154 $\frac{1}{3}\pi$

156 $\theta(x)=0, x\in[-1,1]$.

158 $u(x)=ax, a\in\mathbb{R}, x\in[-1,1]$.

162 $\operatorname{gn}(3-2x)$.

163 30°.

164 $p(x)=\frac{1}{4}x^2$.

166 $\{1,x-1\}$.

169 $\frac{1}{4}e^2+\frac{1}{4}$.

171 $\frac{6}{25}e^{\frac{5}{2}}+\frac{4}{25}$.

173 $\sqrt{\frac{4}{25}e^5+\frac{1}{25}}$.

175 $\frac{2}{9}e^{\frac{3}{2}}+\frac{4}{9}$.

178 $\operatorname{gn}(1+e^2-4x)$.

182 $\sqrt{n+1}$.

184 $\arccos\left(\frac{1}{2}\sqrt{\frac{6n}{2n+1}}\right)$.

186 $\left\{\frac{1}{\sqrt{n+1}}, 2\sqrt{\frac{3n}{(n+2)(n+1)}}\left(x-\frac{1}{2}\right)\right\}$.

193 1.

195 $\sqrt{2}$.

197 $\{1,x-1\}$.

199 $\dfrac{n!}{2^{n+1}}$.

201 $\frac{1}{3}\sqrt{3}$.

203 $p(x)=\frac{3}{4}-\frac{1}{4}x$.

205 $\langle\cos x, e^{-x}\rangle=\frac{2}{5}$ y $\langle\operatorname{sen} x, e^{-x}\rangle=\frac{1}{5}$.

207 $\frac{1}{5}\sqrt{10}$.

209 $\left\{\sqrt{\frac{5}{3}}\cos x, \sqrt{3}(\operatorname{sen} x-\frac{1}{3}\cos x)\right\}$.

211 $Q=\begin{bmatrix}\frac{1}{7}\sqrt{7} & \frac{3}{35}\sqrt{35} & \frac{2}{15}\sqrt{15}\\ \frac{2}{7}\sqrt{7} & -\frac{1}{35}\sqrt{35} & \frac{1}{15}\sqrt{15}\\ -\frac{1}{7}\sqrt{7} & \frac{4}{35}\sqrt{35} & \frac{1}{15}\sqrt{15}\\ \frac{1}{7}\sqrt{7} & \frac{3}{35}\sqrt{35} & -\frac{1}{5}\sqrt{15}\end{bmatrix}, R=\begin{bmatrix}\sqrt{7} & -\frac{1}{7}\sqrt{7} & -\frac{1}{7}\sqrt{7}\\ 0 & \frac{5}{7}\sqrt{35} & \frac{4}{35}\sqrt{35}\\ 0 & 0 & \frac{2}{5}\sqrt{15}\end{bmatrix}$

213 $Q=\begin{bmatrix}\frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} & -\frac{2}{15}\sqrt{15}\\ 0 & \frac{1}{3}\sqrt{3} & \frac{1}{15}\sqrt{15}\\ \frac{1}{3}\sqrt{3} & 0 & \frac{1}{5}\sqrt{15}\\ -\frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} & \frac{1}{15}\sqrt{15}\end{bmatrix}, R=\begin{bmatrix}\sqrt{3} & 0 & 0\\ 0 & \sqrt{3} & -\frac{1}{3}\sqrt{3}\\ 0 & 0 & \frac{1}{3}\sqrt{15}\end{bmatrix}$.

215 Matriz ortogonal.

217 Matriz no ortogonal.

219 Matriz ortogonal.

221 Matriz ortogonal.

234 $\left(\frac{\sqrt{\pi}}{4}\left(1-e^{-1}\right)\right)^{1/2}$.

236 $\frac{1}{4}\sqrt{\pi}e^{-\frac{1}{4}}$.

239 $\frac{1}{8}\sqrt{\pi}e^{-\frac{1}{4}}$.

240 $p^*(x)=\frac{5}{4}e^{-\frac{1}{4}}-\frac{1}{2}e^{-\frac{1}{4}}x^2$.

242 $p^*=\left(\frac{8}{11},\frac{2}{11},\frac{14}{11}\right)$.

244 $p^*=\left(\frac{99}{71},\frac{26}{71},\frac{144}{71},\frac{155}{71}\right)$.

246 $p^*(x)=\frac{1}{2}\ln(3), x\in[1,3]$.

248 $p^*(x)=8e^{-1}-2+(6-18e^{-1})x, x\in[0,1]$.

250 $p^*(x)=\frac{1}{2}e^2-\frac{1}{2}, x\in[0,2]$.

252 $p^*(x)=\pi-2\operatorname{sen} x$.

254 $p^*(x)=-2+4x, x\in[0,\infty]$.

256 $p(x)=-1.9900+2.9920x$.

258 $p(x)=1.0183+1.9968x$.

260 $p(x)=1.0225+1.9425x+1.0175x^2$.

262 $p(x)=0.0135-1.0065x+1.0025x^2$.

264 *(b)* 1 006.7 Ω.

4.2 Espacios vectoriales normados: páginas 400-414

266 2.

268 14.

270 20.

272 7.

274 10.

276 9.

278 2.

280 5.

282 $2\sqrt{2}$.

284 $2\ln 2$.

286 $\dfrac{n(n-1)(2n-1)}{6}$

288 $\dfrac{\sqrt{2}}{\left(\sqrt{2}-1\right)^2+1}$.

290 $\frac{3}{2}\ln 2-\frac{1}{2}$.

292 $\frac{1}{3}$

294 $\frac{37}{6}$.

296 $2\sqrt{2}-2$.

298 $\ln 2$.

301 1.

303 $e-1$.

305 1.

307 $\dfrac{\ln(3)}{3}$.

309 2.

313 $\|f\|=1$.

315 $\|f\|=27\ln 3+9$.

318 $\|f\|=1+e^{-1}$.

319 $\|f\|=54\ln 3+9$.

321 $\|f\|=\frac{1}{2}\sqrt{2}+\frac{1}{8}\pi\sqrt{2}$.

323 $\|f\|=18e^4$.

329 1.

331 $e-1$.

333 2.

335 e^2-1.

337 $\ln(2)$.

341 $(a_n)\notin \ell_2$.

344 $\left(\dfrac{5}{4}\right)^{1/5}$.

346 $\left(\dfrac{1-2^{1-p}}{p-1}\right)^{1/p}$.

348 $\left(\dfrac{1}{4}e^2-\dfrac{1}{4}\right)^{1/2}$.

350 $(10\pi^3-24\pi)^{1/3}$.

352 $(1/4)^{1/3}$.

354 $\left(1-\frac{1}{4}\pi\right)^{1/2}$.

357 $f\notin \mathscr{L}_1\,[1,\infty)$.

359 $\left(\dfrac{a^{1-p}}{p-1}\right)^{1/p}$.

361 1.

363 $(\ln(2))^{-1/2}$.

375 *(b)* $\langle (x_1,x_2),(y_1,y_2)\rangle = 2x_1y_1-x_2y_1-y_2x_1+x_2y_2$.

377 *(b)* $\langle (x_1,x_2),(y_1,y_2)\rangle = 5x_1y_1+3x_2y_1+3y_2x_1+2x_2y_2$.

379 *(b)* $\langle (x_1,x_2,x_3),(y_1,y_2,y_3)\rangle = 5x_1y_1+2x_2y_2-y_2x_1-x_2y_1+x_3y_3-2x_3y_1-2y_3x_1$.

381 *(b)* $\langle (x_1,x_2,x_3),(y_1,y_2,y_3)\rangle = x_1y_1+x_2y_2+2x_3y_3+x_2y_3+y_2x_3$.

383 *(b)* $\langle (x_1,x_2,x_3,x_4),(y_1,y_2,y_3,y_4)\rangle = x_1y_1+2x_2y_2+x_3y_3+x_2y_3+y_2x_3+2x_4y_4-x_2y_4-y_2x_4$.

385 No proviene de un producto interior.

387 No proviene de un producto interior.

389 No proviene de un producto interior.

391 No proviene de un producto interior.

393 Proviene del producto interior $\langle (a_n),(b_n)\rangle = \sum_{n=1}^{\infty}|a_n||b_n|$.

395 No proviene de un producto interior.

397 No proviene de un producto interior.

410
$$\begin{aligned}
B_1(x) &= (e-1)x+1\\
B_2(x) &= \left(1-2e^{\frac{1}{2}}+e\right)x^2+\left(2e^{\frac{1}{2}}-2\right)x+1\\
B_3(x) &= \left(3e^{\frac{1}{3}}-1-3e^{\frac{2}{3}}+e\right)x^3+\left(3+3e^{\frac{2}{3}}-6e^{\frac{1}{3}}\right)x^2+\left(-3+3e^{\frac{1}{3}}\right)x+1\\
B_4(x) &= \left(1+6e^{\frac{1}{2}}-4e^{\frac{1}{4}}-4e^{\frac{3}{4}}+e\right)x^4+\left(-4+4e^{\frac{3}{4}}+12e^{\frac{1}{4}}-12e^{\frac{1}{2}}\right)x^3+\left(6-12e^{\frac{1}{4}}+6e^{\frac{1}{2}}\right)x^2+\left(-4+4e^{\frac{1}{4}}\right)x+1
\end{aligned}$$

412
$$\begin{aligned}
B_1(x) &= (\operatorname{sen} 1)x\\
B_2(x) &= \left(\operatorname{sen} 1-2\operatorname{sen}\frac{1}{2}\right)x^2+2\left(\operatorname{sen}\frac{1}{2}\right)x\\
B_3(x) &= \left(3\operatorname{sen}\frac{1}{3}-3\operatorname{sen}\frac{2}{3}+\operatorname{sen} 1\right)x^3+\left(-6\operatorname{sen}\frac{1}{3}+3\operatorname{sen}\frac{2}{3}\right)x^2+3\left(\operatorname{sen}\frac{1}{3}\right)x
\end{aligned}$$

415 $\vec{p}^* = (0,0)$.

416 $\vec{p}^* = (0,0)$.

417 Todos los vectores de la forma $\vec{p} = a(-1,1), a \in [-1,1]$.

424 $\vec{p}^* = (-1/3, 1/3, 2/3)$.

425 $\vec{p}^* = (0,0,0)$.

426 $\vec{p}^* = (-1/2, 1/2, 1)$.

437 $p^*(x) = 1/2, x \in [0,1]$.

438 $p^*(x) = 1/2, x \in [0,1]$.

439 $p^*(x) = 1/2, x \in [0,1]$.

Capítulo 5

5.1 Transformaciones lineales: páginas 539-551

1 Sí es una transformación lineal.

3 No es una transformación lineal.

5 Sí es una transformación lineal.

7 Sí es una transformación lineal.

9 Sí es una transformación lineal.

11 No es una transformación lineal.

13 No es una transformación lineal.

15 Sí es una transformación lineal.

17 Sí es una transformación lineal.

19 Sí es una transformación lineal.

21 $A = \begin{bmatrix} -5 & 0 & -1 \\ -3 & -7 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \\ -3 & 3 & -1 \end{bmatrix}$.

23 No es una transformación lineal

25 $A = \begin{bmatrix} 2 & -3 \\ -1 & 1 \end{bmatrix}$.

27 No es una transformación lineal.

29 $A = \begin{bmatrix} 1 & 3 & -1 \\ 2 & -1 & 2 \end{bmatrix}$.

31 $T(x,y,z) = (-y-z, 2y+z, -2y+3z-x, -5y+5z-4x)$.

33 $T(x_1,x_2,x_3,x_4) = \left(\frac{3}{2}x_1 - \frac{5}{2}x_2 - \frac{15}{2}x_4 + 7x_3, x_2 + 3x_4 - 2x_3, -\frac{3}{2}x_1 - \frac{1}{2}x_2 + \frac{5}{2}x_4 - 2x_3\right)$.

35 $T(x,y) = (y, 2y)$.

37 $T\left(\begin{bmatrix} a & b \\ d & c \end{bmatrix}\right) = (-c+3d-2b, 2c-3d+3b, 2d-b+a, d-2b+a)$.

39 $T(x,y) = (x\cos\phi - y\,\text{sen}\,\phi, x\,\text{sen}\,\phi + y\cos\phi)$.

41 $T(x,y) = (-x, y)$.

43 $T(x,y,z) = (-x-y+2z, 2x-z, x+y, x+y+2z)$.

45 $T(x,y,z) = (x+2y, x-y, x+y)$.

47 $T(a+bx+cx^2) = bx^3 + ax^2 - 2ax + b$.

49 $T(a\cos x + b\,\text{sen}\,x) = (a-b, a+2b)$.

50 $T\left(\begin{bmatrix} a & b \\ b & c \end{bmatrix}\right) = (a,b,c)$.

52 $T(x_1,x_2,x_3,x_4) = (2x_1 - 2x_4 + x_2, -5x_3 - x_1 + 6x_4 + 2x_2, 5x_3 + 3x_1 - 8x_4 - x_2)$.

54 $T(x,y,z) = (x+y+z, -2z-2x-y, -2z-2x-y)$.

56 $T(a+bx+cx^2+dx^3) = cx^2 + dx^3$

58 $T(x,y,z) = (-2x-y, z+x, z+3x+y)$.

60 $T(a+bx+cx^2+dx^3) = -2b-a+(2b+a)x+dx^3$.

63 Sí es lineal.

65 Sí es lineal.

67 Sí es lineal.

69 No es lineal.

71 No es lineal.

73 No es lineal.

75 Sí es lineal.

77 Sí es lineal.

79 *(b)* $S^\perp = \text{gn}(\vec{e}_{m+1}, \vec{e}_{m+2}, \ldots, \vec{e}_n)$, donde $\{\vec{e}_1, \ldots, \vec{e}_m, \vec{e}_{m+1}, \ldots, \vec{e}_n\}$ es una base de $\mathbb{R}^n$.

82 *(ii)* $A = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$; *(iii)* $\{(1,0),(1,-1)\}$, $\dim(T(\mathbb{R}^2)) = 2$; *(iv)* $\text{Ker}(T) = \{(0,0)\}$;
(v) T es inyectiva; *(vi)* T es suprayectiva.

84 *(ii)* $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$;

(iii) $\{(1,0,0),(1,1,0)\}$, $\dim(T(\mathbb{R}^2) = 2$;

(iv) $\mathrm{Ker}(T) = \{(0,0)\}$;

(v) T es inyectiva;

(vi) T no es suprayectiva.

86 *(ii)* $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$;

(iii) $\{(1,0,1,1),(0,1,0,1),(0,0,-1,0)\}$, $\dim(T(\mathbb{R}^4) = 3$;

(iv) $\mathrm{Ker}(T) = \{(0,-1,1,0)\}$, $\dim(\mathrm{Ker}(T)) = 1$;

(v) T no es inyectiva;

(vi) T no es suprayectiva.

88 *(ii)* $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 3 & -4 \\ 1 & -2 & 1 \\ 3 & -4 & -1 \end{bmatrix}$;

(iii) $\{(1,1,1,3),(-1,3,-2,-4),(1,-4,1,-1)\}$, $\dim(T(\mathbb{R}^3) = 3$;

(iv) $\mathrm{Ker}(T) = \{(0,0,0)\}$;

(v) T es inyectiva;

(vi) T no es suprayectiva.

90 *(ii)* $A = \begin{bmatrix} 1 & -1 & 1 \\ -2 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix}$;

(iii) $\{(1,-2,1),(-1,1,-1),(1,1,2)\}$, $\dim(T(\mathbb{R}^3) = 3$;

(iv) $\mathrm{Ker}(T) = \{(0,0,0)\}$;

(v) T es inyectiva;

(vi) T es suprayectiva.

106 $\dfrac{n(n-1)}{2}$.

107 *(b)* $\mathrm{Ker}(L)$ es el subespacio de las matrices simétricas; una base para este espacio está formada por las matrices de la forma $M_{kl} = [m_{ij}]$, $k,l \in \{1,2,\ldots,n\}$, $k < n$, con

$$m_{ij} = m_{ji} = \begin{cases} 1, & \text{si } i = l \text{ y } j = k \\ 0, & \text{en otro caso} \end{cases}$$

por tanto $\dim(\mathrm{Ker}(T)) = \frac{n(n+1)}{2}$.

111 *(ii)* $T(\mathbf{E}) = \mathrm{gn}(x,x^2,x^3)$, $\mathrm{Ker}(T) = \mathrm{gn}(1)$;

(iii) $\dim(\mathrm{Ker}(T)) = 1$, $\dim(T(\mathbf{E})) = 3$.

113 *(ii)* $T(\mathbf{E}) = \mathrm{gn}(1)$, $\mathrm{Ker}(T) = \mathrm{gn}(x,x^2,x^3,\ldots)$;

(iii) $\dim(\mathrm{Ker}(T)) = \infty$, $\dim(T(\mathbf{E})) = 1$.

115 *(ii)* $T(\mathbf{E}) = \mathrm{gn}(1)$, $\mathrm{Ker}(T) = \mathrm{gn}\left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right)$;

(iii) $\dim(\mathrm{Ker}(T)) = 3$, $\dim(T(\mathbf{E})) = 1$.

117 *(ii)* $T(\mathbf{E}) = \mathrm{gn}(x,x^2+1,x^3+2x)$, $\mathrm{Ker}(T) = \{0\}$;

(iii) $\dim(\mathrm{Ker}(T)) = 0$, $\dim(T(\mathbf{E})) = 3$.

118 *(ii)* $\mathrm{Ker}(T) = \mathrm{gn}(e^{-x^2/2})$;

(iii) $\dim(\mathrm{Ker}(T)) = 1$, $\dim(T(\mathbf{E})) = \infty$.

122 *(ii)* $T(\mathbf{E}) = \mathrm{gn}(\cos x, \operatorname{sen} x)$;

(iii) $\dim(\mathrm{Ker}(T)) = \infty$, $\dim(T(\mathbf{E})) = 2$.

124 *(ii)* $\mathrm{Ker}(T) = \mathrm{gn}((u_n))$, donde $u_n = 1\ \forall n$;

(iii) $\dim(\mathrm{Ker}(T)) = 1$, $\dim(T(\mathbf{E})) = \infty$.

129 *(a)* $T(\mathbf{E}) = \mathrm{gn}(1,\cos x,\operatorname{sen} x)$, $\dim(T(\mathbf{E})) = 3$.

(b) $\mathrm{Ker}(T) = S$.

(c) *(i)* Todas las funciones $f \in S - \{0\}$ para $\lambda = 0$. *(ii)* Las funciones $f(x) = b\cos x + c\operatorname{sen} x$, con uno de los dos coeficientes distinto de cero, para $\lambda = \pi$; y las funciones $f(x) = a \neq 0$ para $\lambda = 2\pi$.

131 $T^{-1}(x,y) = (-x+2y, x-y)$.

133 $T^{-1}(x,y) = (3x-y, 2x-y)$.

135 $T^{-1}(x,y,z) = (\frac{1}{2}x+\frac{1}{4}y-\frac{1}{4}z, \frac{1}{2}x-\frac{1}{4}y+\frac{1}{4}z, -\frac{1}{2}y-\frac{1}{2}z)$.

137 $T(x,y,z) = (2x-y+z, -x+y-z, -5x+3y-2z)$.

149 *(a)* *(i)* $T_1T_2(x,y,z) = (x, x+y+z, x+y)$. *(ii)* $T_2T_1(x,y,z) = (x, x+z, x+y+z)$. *(iii)* $T_1^2(x,y,z) = (x,y,z)$. *(iv)* $T_2^2(x,y,z) = (x, 2x+y, 3x+2y+z)$. *(v)* $(T_1T_2 - T_2T_1)^2(x,y,z) = (0,y,z)$.

(b) *(i)* $T_1^{-1}(x,y,z) = (x,z,y)$. *(ii)* $T_2(x,y,z) = (x, -x+y, -y+z)$.

(iii) $(T_1T_2)^{-1}(x,y,z) = (x, -x+z, y-z)$. *(iv)* $(T_2T_1)^{-1}(x,y,z) = (x, -y+z, -x+y)$.

(c) $T_2 - I = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}$, $(T_2 - I)^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $(T_2 - I)^n = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ si $n \geq 3$.

155 (a) (i) $D(p(x)) = -2 + 6x + 3x^2$, (ii) $T(p(x)) = -2x + 6x^2 + 3x^3$, (iii) $DT(p(x)) = -2 + 12x + 9x^2$, (iv) $TD(p(x)) = 6x + 6x^2$, (v) $(DT - TD)(p(x)) = -2 + 6x + 3x^2$. (vi) $(T^2D^2 - D^2T^2)(p(x)) - 24 + 48x$.

(b) $p(x) = ax$, $a \in \mathbb{R}$.

(c) $p(x) = a_0 + a_2x^2$, $a_0, a_2 \in \mathbb{R}$.

(d) $p \in \mathbf{P}$.

170 $\nu = 4$.

5.2 Representación matricial: páginas 552-568

171 $[\vec{u}]_{\mathscr{B}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$.

173 $[\vec{u}]_{\mathscr{B}} = \begin{bmatrix} \frac{1}{2} \\ -4 \end{bmatrix}$.

175 $[\vec{u}]_{\mathscr{B}} = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$.

177 $[\vec{u}]_{\mathscr{B}} = \begin{bmatrix} -13 \\ 0 \\ 8 \end{bmatrix}$.

179 $[\vec{u}]_{\mathscr{B}} = \begin{bmatrix} -25 \\ -6 \\ -6 \\ 7 \end{bmatrix}$.

181 $[\vec{u}]_{\mathscr{B}} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$.

183 $[\vec{u}]_{\mathscr{B}} = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}$.

185 $[\vec{u}]_{\mathscr{B}} = \begin{bmatrix} \frac{7}{5} \\ \frac{4}{5} \\ -\frac{6}{5} \end{bmatrix}$.

187 $P = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$.

189 $P = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -2 \\ 3 & -3 & 5 \end{bmatrix}$.

191 $P = \begin{bmatrix} 1 & -2 & 1 & 3 \\ -2 & 5 & -2 & -5 \\ 3 & -6 & 4 & 9 \\ 4 & -8 & 4 & 11 \end{bmatrix}$.

193 $P = \begin{bmatrix} -1 & -2 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$.

195 $P = \begin{bmatrix} 2 & -1 & -2 & -1 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

197 (i) $P = \begin{bmatrix} 1 & 2 \\ -3 & -5 \end{bmatrix}$,

(ii) $P^{-1} = \begin{bmatrix} -5 & -2 \\ 3 & 1 \end{bmatrix}$.

199 (i) $P = \begin{bmatrix} 5 & 9 \\ 22 & 37 \end{bmatrix}^{-1}$,

(ii) $P^{-1} \begin{bmatrix} -\frac{37}{13} & \frac{9}{13} \\ \frac{22}{13} & -\frac{5}{13} \end{bmatrix}$.

201 (i) $P = \begin{bmatrix} -4 & 0 & -3 \\ 3 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}$,

(ii) $P^{-1} = \begin{bmatrix} 2 & 3 & 0 \\ 3 & 4 & 1 \\ -3 & -4 & 0 \end{bmatrix}$.

203 (i) $P = \begin{bmatrix} -3 & 4 & -4 \\ -2 & -3 & 1 \\ 1 & -4 & 3 \end{bmatrix}$,

(ii) $P^{-1} = \begin{bmatrix} 5 & -4 & 8 \\ -7 & 5 & -11 \\ -11 & 8 & -17 \end{bmatrix}$.

205 (i) $P = \begin{bmatrix} 1 & -4 & 1 \\ -3 & 5 & 0 \\ 4 & -4 & -1 \end{bmatrix}$,

(ii) $P^{-1} = \begin{bmatrix} 5 & 8 & 5 \\ 3 & 5 & 3 \\ 8 & 12 & 7 \end{bmatrix}$.

207 (i) $P = I_3$,

(ii) $P^{-1} = I_3$.

210 (ii) $[T]_{\mathscr{B}} = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$,

(iii) $[T(\vec{u})]_{\mathscr{B}} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$.

212 (ii) $[T]_{\mathscr{B}} = \begin{bmatrix} 3 & 2 \\ -1 & -3 \end{bmatrix}$,

(iii) $[T(\vec{u})]_{\mathscr{B}} = \begin{bmatrix} -11 \\ 34 \end{bmatrix}$.

214 *(ii)* $[T]_{\mathscr{B}} = \begin{bmatrix} -8 & 0 & -20 \\ 7 & 2 & 14 \\ 1 & -5 & 11 \end{bmatrix}$,

(iii) $[T(u)]_{\mathscr{B}} = \begin{bmatrix} 28 \\ -19 \\ -15 \end{bmatrix}$.

216 *(ii)* $[T]_{\mathscr{B}} = \begin{bmatrix} -4 & 3 & -7 \\ 3 & -2 & 5 \\ 4 & -3 & 7 \end{bmatrix}$,

(iii) $[T(\vec{u})]_{\mathscr{B}} = \begin{bmatrix} -5 \\ 3 \\ 5 \end{bmatrix}$.

218 *(i)* $[T]_{\mathscr{B}} = \begin{bmatrix} 10 & -21 & 8 & 23 \\ 0 & 0 & -1 & -1 \\ -9 & 19 & -9 & -25 \\ 0 & 0 & 0 & 1 \end{bmatrix}$,

(iii) $[T(\vec{u})]_{\mathscr{B}} = \begin{bmatrix} 14 \\ 1 \\ 2 \\ -5 \end{bmatrix}$.

219 *(a)* $T(2\hat{\imath}-3\hat{\jmath}) = -8\hat{\imath}-\hat{\jmath}$, $T^2(2\hat{\imath}-3\hat{\jmath}) = 6\hat{\imath}-9\hat{\jmath}$.

(b) $T(x\hat{\imath}+y\hat{\jmath}) = (x-2y)\hat{\imath}+(x+y)\hat{\jmath}$.

(c) $\mathrm{Ker}(T) = \{(0,0)\}$, $T(\mathbb{R}^2) = \mathrm{gn}(-\hat{\imath}+\hat{\jmath}, 2\hat{\imath}+\hat{\jmath})$; $\dim(\mathrm{Ker}(T)) = 0$, $\dim(T(\mathbb{R}^2)) = 2$.

(d) $[T]_{\mathscr{B}} = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}$.

(e) $[T]_{\mathscr{B}} = \begin{bmatrix} -\frac{9}{4} & \frac{11}{4} \\ -\frac{3}{4} & \frac{9}{4} \end{bmatrix}$.

(f) $[T^2]_{\mathscr{B}} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$.

(g) $[T^2]_{\mathscr{B}} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$.

221 *(ii)* $[T]_{\mathscr{B}} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{4}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

223 *(ii)* $[T]_{\mathscr{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

225 *(ii)* $[T]_{\mathscr{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$.

227 *(ii)* $[T]_{\mathscr{B}} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 0 & 1 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$.

229 *(ii)* $[T]_{\mathscr{B}} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$.

231 *(ii)* $[T]_{\mathscr{B}} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$.

233 $[D]_{\mathscr{B}} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

235 $[D]_{\mathscr{B}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

237 $[D]_{\mathscr{B}} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$.

239 $[T]_{\mathscr{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$.

240 $[D]_{\mathscr{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

241 $[DT]_{\mathscr{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

242 $[TD]_{\mathscr{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

243 $[TD-DT]_{\mathscr{B}} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

244 $[T^2]_{\mathscr{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}$.

245 $[T^2D^2 - D^2T^2]_{\mathscr{B}} = \begin{bmatrix} 0 & 0 & -8 & 0 \\ 0 & 0 & 0 & -48 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

247 *(i)* $P = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$.

(ii) $[T]_{\mathscr{B}} = \begin{bmatrix} 23 & 15 \\ -35 & -23 \end{bmatrix}$.

249 *(i)* $P = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$.

(ii) $[T]_{\mathscr{B}} = \begin{bmatrix} 6 & 10 & 32 \\ 3 & 7 & 19 \\ -2 & -4 & -11 \end{bmatrix}$.

251 *(i)* $P = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$.

(ii) $[T]_{\mathscr{B}} = \begin{bmatrix} -4 & -11 & -2 \\ 2 & 4 & 2 \\ 3 & 4 & 4 \end{bmatrix}$

253 *(i)* $P = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & -1 & 2 & 2 \\ 1 & -1 & 2 & 1 \\ 3 & -3 & 3 & 2 \end{bmatrix}$.

(ii) $[T]_{\mathscr{B}} = \begin{bmatrix} 13 & -13 & 13 & 8 \\ 6 & -6 & 7 & 4 \\ 0 & 0 & 1 & 0 \\ -9 & 9 & -9 & -5 \end{bmatrix}$.

255 *(i)* $P = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 2 & -1 & 2 & -2 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & -1 & 2 \end{bmatrix}$.

(ii) $[T]_{\mathscr{B}} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & -1 \\ 3 & -2 & 4 & -3 \\ 2 & -1 & 3 & -1 \end{bmatrix}$.

257 *(a)* $P = \begin{bmatrix} 2 & 1 & -1 & -3 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(b) $[T]_{\mathscr{B}_1} = \begin{bmatrix} -\frac{1}{3} & \frac{1}{3} & -\frac{2}{3} & 1 \\ -\frac{1}{3} & \frac{1}{3} & \frac{4}{3} & 1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

259 *(ii)* $\det(T) = 3$;
(iii) $T^{-1}(x,y) = (\frac{1}{3}x + \frac{2}{3}y, -\frac{1}{3}x + \frac{1}{3}y)$.

261 *(ii)* $\det(T) = 0$;
(iii) T no es invertible.

263 *(ii)* $\det(T) = 1$;
(iii) $T^{-1}(x,y,z) = (32x + 19y - 12z, 8x + 5y - 3z, -5x - 3y + 2z)$

265 *(ii)* $\det(T) = 0$;
(iii) T no es invertible.

267 *(ii)* $\det(T) = 0$;
(iii) T no es invertible.

269 *(ii)* $\det(T) = 0$;
(iii) T no es invertible.

271 *(ii)* $\det(T) = 1$;
(iii) $T^{-1}(p) = q$, donde $q(x) = p(x-1)$.

280 *(iii)* $[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} -8 & 2 & -6 \\ 11 & -2 & 7 \end{bmatrix}$.

282 *(iii)* $[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} 5 & 1 \\ 4 & 2 \\ -1 & -1 \end{bmatrix}$.

284 *(iii)* $[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} 5 & 3 \\ 5 & 0 \\ -7 & -2 \end{bmatrix}$.

286 *(iii)* $[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} -1 & -9 & 2 & -1 \\ 3 & 8 & -5 & 0 \\ 1 & -2 & -2 & -1 \end{bmatrix}$.

288 *(iii)* $[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} 6 & 4 \\ -9 & 23 \\ -2 & 12 \\ 0 & -10 \end{bmatrix}$.

290 *(iii)* $[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \end{bmatrix}$.

292 *(iii)* $[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} 0 & 1 & 0 \\ 3 & 3 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

294 *(iii)* $[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$.

296 $[T(\vec{u})]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

298 $[T(\vec{u})]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} -5 \\ -1 \\ -5 \end{bmatrix}$.

300 $[T(\vec{u})]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} -8 \\ -4 \\ 1 \end{bmatrix}$.

302 $[T(\vec{u})]_{\mathscr{B}_1}^{\mathscr{B}_2} = [2]$.

303 *(i)* $P = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 1 \end{bmatrix}$.

(ii) $Q = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. *(iii)* $[T]_{\mathscr{B}'_1}^{\mathscr{B}'_2} = \begin{bmatrix} 2 & 0 & 4 \\ -4 & 1 & -6 \end{bmatrix}$.

305 *(i)* $P=\begin{bmatrix}1&0&-1&0\\0&1&1&-1\\-1&1&1&0\\1&1&0&0\end{bmatrix}$. *(ii)* $Q=\begin{bmatrix}1&1&-1\\6&5&-6\\1&-1&2\end{bmatrix}$.

(iii) $[T]_{\mathscr{B}_1'}^{\mathscr{B}_2'}=\begin{bmatrix}-\frac{1}{3}&\frac{1}{3}&1&-\frac{1}{3}\\4&1&-4&0\\\frac{8}{3}&\frac{4}{3}&-2&-\frac{1}{3}\end{bmatrix}$.

310 *(a)* $T(x,y)=(0,x-y,x+y)$.

(b) $\dim(\mathrm{Ker}(T))=0$, $\dim(T(\mathbb{R}^2))=2$, $\mathrm{gn}(T(\mathbb{R}^2))=((0,1,1),(0,-1,1))$.

(c) $[T]=\begin{bmatrix}0&0\\1&-1\\1&1\end{bmatrix}$.

(d) $\mathscr{B}_1=((1,0),(0,1))$, $\mathscr{B}_2=((0,1,1),(0,-1,1),(1,0,0))$.

312 $\mathbf{F}=\mathrm{gn}(x^2,x)$; $\mathscr{B}_1=\{\frac{1}{6}x^3,\frac{1}{2}x^2,x,1\}$, $\mathscr{B}_2=\{x^2,x\}$.

314 $f(x,y)=-\frac{1}{4}x-\frac{13}{4}y$; $f(-2,1)=-\frac{11}{4}$.

315 Sí.

317 Sí.

319 Sí.

321 Sí.

323 Sí.

325 No.

326 $\mathscr{B}^*=\{\phi_1,\phi_2\}$, donde $\phi_1(x,y)=-x+2y$ y $\phi_2(x,y)=x-y$, $(x,y)\in\mathbb{R}^2$.

328 $\mathscr{B}^*=\{\phi_1,\phi_2,\phi_3\}$, donde $\phi_1(x,y,z)=y-4x+z$, $\phi_2(x,y,z)=y-2x$ y $\phi_1(x,y,z)=-z+3x$, $(x,y,z)\in\mathbb{R}^3$.

330 $\mathscr{B}^*=\{\phi_1,\phi_2,\phi_3\}$, donde $\phi_i(a_0+a_1x+a_2x^2)=a_i$, $i=0,1,2$; $a_0+a_1x+a_2x^2\in\mathbf{P}_2$.

332 $\mathscr{B}^*=\{\phi_1,\phi_2,\phi_3\}$, donde $\phi_1(a_0+a_1x+a_2x^2)=\frac{1}{3}a_0-\frac{2}{3}a_1+\frac{1}{3}a_2,$; $\phi_2(a_0+a_1x+a_2x^2)=\frac{1}{3}a_1+\frac{1}{3}a_0+\frac{1}{3}a_2$, $\phi_3(a_0+a_1x+a_2x^2)=a_2$; $a_0+a_1x+a_2x^2\in\mathbf{P}_2$.

335 *(ii)* $\mathscr{B}=\{\vec{e}_1,\ldots,\vec{e}_n\}$, la base canónica de $\mathbb{R}^n$.

337 *(ii)* $\mathscr{B}=\{(3,-2),(-4,3)\}$.

339 *(ii)* $\mathscr{B}=\{(0,-1,1),(-1,1,0),(-1,-1,1)\}$.

341 *(ii)* $\mathscr{B}=\{-2+3x,3-4x\}$.

343 *(ii)* $\mathscr{B}=\{3x-\frac{3}{2}x^2,1-3x+\frac{3}{2}x^2,-\frac{1}{2}x+\frac{3}{4}x^2\}$.

345 *(ii)* $\mathscr{B}=\left\{\begin{bmatrix}-3&-2\\-1&-3\end{bmatrix},\begin{bmatrix}1&1\\0&0\end{bmatrix},\begin{bmatrix}-1&0\\1&0\end{bmatrix},\begin{bmatrix}-1&0\\0&-1\end{bmatrix}\right\}$.

346 $n-1$.

349 $\vec{u}_f=f(1,0)(1,0)+f(0,1)(0,1)=(2,-3)$.

351 $\vec{u}_f=f(1,0,0)(1,0,0)+f(0,1,0)(0,1,0)+f(0,0,1)(0,0,1)=(1,-1,1)$

353 $\vec{u}_f=\begin{bmatrix}1&-2\\3&1\end{bmatrix}$

355 $\vec{u}_f=\frac{1}{2}+\frac{3}{4}x$.

360 $T^*(x,y)=(x+2y,x-5y)$.

362 $T^*(a,b,c,d)=(a+c+d,-a-2c,2b,b+3d)$.

364 $T^*(a+bx+cx^2)=(27c+30b+30a)x^2+(-26b-24c-24a)x+3b+3c+2a$.

375 $S=\mathrm{gn}(0,1,1,0)$; $\{\varphi_1,\varphi_2,\varphi_3\}$ es una base de S^0, donde $\varphi_1(x_1,x_2,x_3,x_4)=x_1$, $\varphi_2(x_1,x_2,x_3,x_4)=x_2-x_3$ y $\varphi_3(x_1,x_2,x_3,x_4)=x_4$; $\dim(S^0)=3$. También $\{f_1,f_2,f_3\}$ es una base de S^0.

376 Una base de V^0 es $\{\varphi_1,\varphi_2\}$, donde $\varphi_1(x_1,x_2,x_3,x_4)=-4x_1+x_2+x_4$ y $\varphi_2(x_1,x_2,x_3,x_4)=3x_1-x_2+x_3$, $(x_1,x_2,x_3,x_4)\in\mathbb{R}^4$.

377 *(a)* $T^t(\phi)=\varphi$, donde $\varphi(x,y,z)=2x-3y+4z$, para cada $(x,y,z)\in\mathbb{R}^3$.

378 *(b)* $T^t(f) = \varphi$, donde $\varphi(a_0 + a_1x) = -2a_1 - 2a_0$. *(c)* $[T^t]_{\mathscr{B}_1^*}^{\mathscr{B}_2^*} = \begin{bmatrix} 0 & 2 \\ -1 & 1 \end{bmatrix}$. *(d)* $[T]_{\mathscr{B}_1}^{\mathscr{B}_2} = \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix}$.

5.3 Valores y vectores propios, diagonalización: páginas 568-578

403 $\lambda \in \mathbb{R}$, con funciones propias correspondientes $\varphi(x) = Cx^\lambda$, $C \in \mathbb{R} - \{0\}$, $x \in (0, \infty)$.

405 *(c)* $\lambda < 0$ con funciones propias correspondientes $\varphi(x) = Ce^{\frac{x^2}{2\lambda}}$, $C \in \mathbb{R} - \{0\}$.

407 *(c)* $\lambda = -n^2$, con funciones propias correspondientes $\varphi(x) = C\operatorname{sen}(nx)$, $C \in \mathbb{R} - \{0\}$, $n = 1, 2, \ldots$

409 $\lambda = 1$, con vectores propios correspondientes las matrices no nulas de la forma $\begin{bmatrix} a & b \\ b & d \end{bmatrix}$ (las matrices simétricas no nulas); $\lambda = 0$, con vectores propios correspondientes las matrices no nulas de la forma $\begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$, $b \neq 0$ (las matrices antisimétricas no nulas).

421 No.

423 *(i)* $\lambda_1 = \lambda_2 = -1$;
(ii) $E_{-1} = \text{gn}((0,1))$;
(iii) $\dim(E_{-1}) = 1$.

425 *(i)* $\lambda_1 = 1$, $\lambda_2 = -1$;
(ii) $E_1 = \text{gn}((2,1))$, $E_{-1} = \text{gn}((3,1))$;
(iii) $\dim(E_1) = 1$, $\dim(E_{-1}) = 1$.

427 *(i)* $\lambda_1 = 1$, $\lambda_2 = 2$;
(ii) $E_1 = \text{gn}((-1,1))$, $E_2 = \text{gn}((1,-2))$;
(iii) $\dim(E_1) = 1$, $\dim(E_2) = 1$.

429 *(i)* $\lambda_1 = \frac{3}{2} + \frac{1}{2}\sqrt{33}$, $\lambda_2 = \frac{3}{2} - \frac{1}{2}\sqrt{33}$;
(ii) $E_{\lambda_1} = \text{gn}((1, \frac{5}{2} + \frac{1}{2}\sqrt{33}))$, $E_2 = \text{gn}((1, \frac{5}{2} - \frac{1}{2}\sqrt{33}))$;
(iii) $\dim(E_{\lambda_1}) = 1$, $\dim(E_{\lambda_2}) = 1$.

431 *(i)* $\lambda_1 = -1$, $\lambda_2 = 1 = \lambda_3$;
(ii) $E_{-1} = \text{gn}((-1,-1,1))$, $E_1 = \text{gn}((1,2,0),(0,0,1))$;
(iii) $\dim(E_{-1}) = 1$, $\dim(E_1) = 2$.

433 *(i)* $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$;
(ii) $E_1 = \text{gn}((1,0,0))$, $E_2 = \text{gn}((1,1,1))$, $E_3 = \text{gn}((-1,2,1)$;
(iii) $\dim(E_1) = 1$, $\dim(E_2) = 1$, $\dim(E_3) = 1$.

435 *(i)* $\lambda_1 = -1 = \lambda_2$, $\lambda_3 = 2$;
(ii) $E_{-1} = \text{gn}((1,0,2),(0,1,0))$, $E_2 = \text{gn}((1,1,1))$;
(iii) $\dim(E_{-1}) = 2$, $\dim(E_2) = 1$.

437 *(i)* $\lambda_1 = 2$, $\lambda_2 = \lambda_3 = 3$;
(ii) $E_2 = \text{gn}((-2,0,1))$, $E_3 = \text{gn}((-3,-1,1))$;
(iii) $\dim(E_2) = 1$, $\dim(E_3) = 1$.

439 *(i)* $\lambda_1 = 1$, $\lambda_2 = -2 + \sqrt{2}$, $\lambda_3 = -2 - \sqrt{2}$;
(ii) $E_{\lambda_1} = \text{gn}((1,0,0))$, $E_{\lambda_2} = \text{gn}((-11/7 + \sqrt{2}/7, \sqrt{2} - 1, 1)$,
$E_{\lambda_3} = \text{gn}((-11/7 - \sqrt{2}/7, -\sqrt{2} - 1, 1))$; $\dim(E_{\lambda_1}) = 1$, $\dim(E_{\lambda_2}) = 1$, $\dim(E_{\lambda_3}) = 1$.

442 Matriz diagonalizable: $C = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$, $D = \begin{bmatrix} -3 & 0 \\ 0 & 1 \end{bmatrix}$.

444 Matriz diagonalizable: $C = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

446 Matriz no diagonalizable.

448 Matriz no diagonalizable.

450 Matriz no diagonalizable.

452 Matriz diagonalizable: $C = \begin{bmatrix} 8 & 5 & 3 \\ 5 & 3 & 2 \\ 1 & 0 & 0 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

454 Matriz diagonalizable: $C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

456 Matriz diagonalizable: $C = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{7}{5} & 1 & 1 \\ -\frac{3}{5} & 1+\frac{\sqrt{6}}{2} & 1-\frac{\sqrt{6}}{2} \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{6} & 0 \\ 0 & 0 & -\sqrt{6} \end{bmatrix}$.

458 Matriz diagonalizable: $C = \begin{bmatrix} 2+\frac{\sqrt{2}}{2} & 2-\frac{\sqrt{2}}{2} & 0 \\ 1 & 1 & 0 \\ \frac{10}{7}+\frac{9}{14}\sqrt{2} & \frac{10}{7}-\frac{9}{14}\sqrt{2} & 1 \end{bmatrix}$, $D = \begin{bmatrix} -2+\sqrt{2} & 0 & 0 \\ 0 & -2-\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

460 Matriz no diagonalizable.

462 Matriz diagonalizable: $C = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$, $D = \begin{bmatrix} -3 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

464 *(b)* $\lambda_k = k$, $k = 1, 2, \ldots, n+1$.

(c) T es diagonalizable: $C = I_{n+1}$, $D = \operatorname{diag}(1, 2, \ldots, n+1)$.

466 *(ii)* Diagonalizable, $C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

469 No, pues no es un múltiplo de la identidad.

470 Que $(a-d)^2 + 4bc > 0$ o $a = d$ y $b = c = 0$.

472 $a = c = 0$.

473 *(i)* $\lambda_1 = -2$, $\lambda_2 = 2$;

(ii) $\mathscr{B}_{\lambda_1} = \{(5, -2-i)\}$, $\mathscr{B}_{\lambda_2} = \{(5, 2-i)\}$;

(iii) matriz diagonalizable: $C = \begin{bmatrix} 5 & 5 \\ -2-i & 2-i \end{bmatrix}$, $D = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$.

475 *(i)* $\lambda_1 = 1 + \sqrt{2}i$, $\lambda_2 = 1 - \sqrt{2}i$,

(ii) $\mathscr{B}_{\lambda_1} = \{(-\sqrt{2}i, 1)\}$, $\mathscr{B}_{\lambda_2} = \{(\sqrt{2}i, 1)\}$;

(iii) matriz diagonalizable: $C = \begin{bmatrix} -\sqrt{2}i & \sqrt{2}i \\ 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 1+\sqrt{2}i & 0 \\ 0 & 1-\sqrt{2}i \end{bmatrix}$.

477 *(i)* $\lambda_1 = \lambda_2 = 2 + i$;

(ii) $\mathscr{B}_{2+i} = \{(1, 0)\}$;

(iii) matriz no diagonalizable.

479 *(i)* $\lambda_1 = 1$, $\lambda_2 = 1 + i$, $\lambda_3 = 1 - i$;

(ii) $\mathscr{B}_{\lambda_1} = \{(0, 1, 1)\}$, $\mathscr{B}_{\lambda_2} = \{(1, 0, -i)\}$, $\mathscr{B}_{\lambda_2} = \{(1, 0, i)\}$;

(iii) matriz diagonalizable: $C = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & -i & i \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1+i & 0 \\ 0 & 0 & 1-i \end{bmatrix}$.

481 *(i)* $\lambda_1 = 3$, $\lambda_2 = \lambda_3 = i$;

(ii) $\mathscr{B}_{\lambda_1} = \{(1, 1, -2)$, $\mathscr{B}_i = \{(1, 0, 0)\}$;

(iii) matriz no diagonalizable.

483 *(i)* $\lambda_1 = 2$, $\lambda_2 = \lambda_3 = 1 - i$;

(ii) $\mathscr{B}_{\lambda_1} = \{(1,-2,1)\}$, $\mathscr{B}_{1-i} = \{(0,0,1),(-1,1,0)\}$;

(iii) matriz diagonalizable: $C = \begin{bmatrix} 1 & 0 & -1 \\ -2 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1-i & 0 \\ 0 & 0 & 1-i \end{bmatrix}$.

486 $A^n = \begin{bmatrix} 2-2^n & -2^n+(-1)^n & -1+(-1)^n \\ -2+2^{n+1} & 2^{n+1}-(-1)^n & 1-(-1)^n \\ 2-2^{n+1} & -2^{n+1}+2(-1)^n & -1+2(-1)^n \end{bmatrix}$, $\det(A) = -2$; $\text{tra}(A) = 2$.

488 $A^n = \begin{bmatrix} 1 & 0 & -1+2(-i)^n-i^n & -(-i)^n+i^n \\ 0 & 1 & -3+2(-i)^n+i^n & 2-(-i)^n-i^n \\ 0 & 0 & 2(-i)^n-i^n & -(-i)^n+i^n \\ 0 & 0 & 2(-i)^n-2i^n & -(-i)^n+2i^n \end{bmatrix}$, $\det(A) = 1$, $\text{tra}(A) = 2$.

491 Si $\{\vec{\xi}_1,\ldots,\vec{\xi}_n\}$ una base ortonormal del espacio $\mathbf{E}$, $\vec{u}_f = \sum_{k=1}^{n} \overline{f(\vec{\xi}_k)}\vec{\xi}_k$.

492 $\vec{u}_0 = (2+i,-1,2-i)$.

502 $S = \text{gn}(x^3,x)$.

505 $p_A = \mathfrak{m}_A = \lambda^2 - 2\lambda + 1$.

507 $p_A = \mathfrak{m}_A = \lambda^2 - 2\lambda$.

509 $p_A = \lambda^3 - 3\lambda - 2$, $\mathfrak{m}_A = \lambda^2 - \lambda - 2$.

511 $p_A = \lambda^4 - 6\lambda^3 + 13\lambda^2 - 12\lambda + 4$, $\mathfrak{m}_A = \lambda^2 - 3\lambda + 2$.

513 $p_A = \lambda^4 + 2\lambda^2 + 1$, $\mathfrak{m}_A = \lambda^2 + 1$.

Capítulo 6

6.1 Matrices de incidencia y teoría de grafos:
páginas 728-730

2 $A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$

4 $A = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

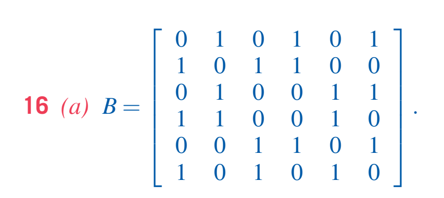
6 $(A^3)_{31} = 3$: $P_3 \to P_4 \to P_2 \to P_1$, $P_3 \to P_1 \to P_4 \to P_1$, $P_3 \to P_1 \to P_2 \to P_1$.

8 $(A^3)_{14} = 1$: $P_1 \to P_2 \to P_3 \to P_4$.

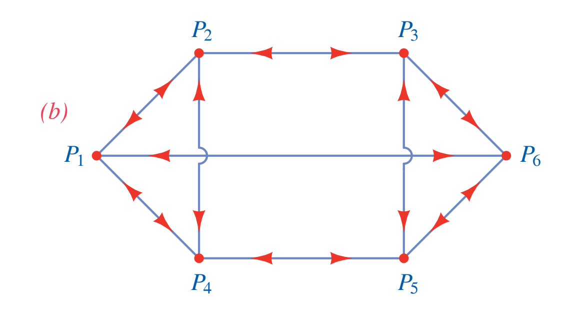
10 $(A^2)_{52} = 1$: $P_5 \to P_3 \to P_2$.

12 $(A + A^2 + A^3)_{43} = 0$

14 La matriz asociada es $B = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ y $B^3 = \begin{bmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 3 & 1 \\ 1 & 2 & 0 & 0 & 1 \\ 2 & 3 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$; puesto que $(B^3)_{ii} = 0$ para todo i, P_i no pertenece a clan alguno para todo $i = 1,2,3$.

16 *(a)* $B =$ 

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

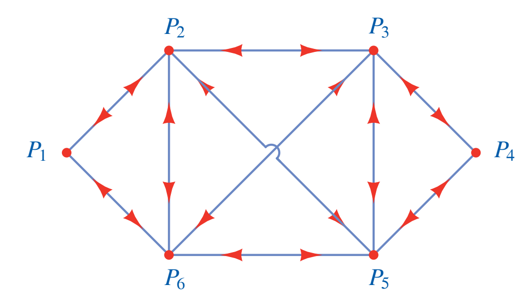


(c) $B = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}^3 = \begin{bmatrix} 2 & 6 & 3 & 6 & 3 & 7 \\ 6 & 2 & 7 & 6 & 3 & 3 \\ 3 & 7 & 2 & 3 & 6 & 6 \\ 6 & 6 & 3 & 2 & 7 & 3 \\ 3 & 3 & 6 & 7 & 2 & 6 \\ 7 & 3 & 6 & 3 & 6 & 2 \end{bmatrix}$. Ya que $(B^3)_{ii} > 0$ para todo i, todo punto P_i es miembro de algún clan.

(d) $\{P_1, P_2, P_4\}$, $\{P_3, P_5, P_6\}$.

18 *(a)* $B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$.

(b)



(c) $B^3 = \begin{bmatrix} 2 & 7 & 4 & 4 & 4 & 7 \\ 7 & 8 & 11 & 4 & 11 & 9 \\ 4 & 11 & 8 & 7 & 9 & 11 \\ 4 & 4 & 7 & 2 & 7 & 4 \\ 4 & 11 & 9 & 7 & 8 & 11 \\ 7 & 9 & 11 & 4 & 11 & 8 \end{bmatrix}$; y ya que $(B^3)_{ii} > 0$ para todo i, todo punto P_i pertenece a un clan.

(d) $\{P_1, P_2, P_6\}$, $\{P_2, P_3, P_5, P_6\}$ y $\{P_3, P_4, P_5\}$.

20 *(b)* Todo individuo, excepto P_4, domina a los demás en a lo más dos etapas; todo individuo es dominado por los demás, excepto por P_4, en a lo más dos etapas.

22 3.

24 P_1 domina a todos los demás, excepto a P_5, en a lo más dos etapas; P_2 únicamente domina a P_3, P_4 y P_6 en a lo sumo dos etapas; P_3 domina a P_4 y P_6 en a lo más dos etapas; P_4 y P_6 no ejercen dominio en individuo alguno; P_5 únicamente domina a P_4 y P_6 en a lo más dos etapas; P_1 y P_5 no son dominados por individuo alguno; P_2 es dominado por P_1 en a lo más dos etapas; P_3 es dominado por P_1 y P_2 en a lo sumo dos etapas; P_4 es dominado por todos los demás, excepto por P_6, en a lo más dos etapas; P_6 es dominado por todos los demás, excepto por P_4, en a lo más dos etapas.

6.2 Redes de conducción y principios de conservación: páginas 730-733

25 $(x_1, x_2, x_3, x_4) = (350, 300, 400, 500)$

27 *(a)* $(f_1, f_2, f_3, f_4, f_5) = (70 - s + r, s - r - 10, 80 - s, s, r); s, r \in \mathbb{R}$.

(b) f_1, f_2, f_3, f_4, f_5 como en *(a)* y r, s los puntos en el polígono limitado por las rectas $-r + s = 70$, $-r + s = 10$, $r = 80$ y $r = 0$ en el plano r, s.

30 $I_1 = 0.857$ A, $I_2 = 1.486$ A, $I_3 = 0.629$ A.

32 $I_1 = 3$ A, $I_2 = 2$ A, $I_3 = -1$ A.

34 $I_1 = 2$ A, $I_2 = 6$ A, $I_3 = 4$ A.

36 $I_1 = 3$ A, $I_2 = 4$ A, $I_3 = 1$ A, $I_4 = 2$ A, $I_5 = 5$ A.

39 $E = 10$ V, $I_2 = 0$ A, $I_3 = -2$ A, $I_4 = -2$ A, $I_5 = 0$ A, $I_6 = 2$ A.

40 $2Na + 2H_2O \longrightarrow 2NaOH + H_2$

42 $2MnO_2 + 4KOH + O_2 \longrightarrow 2K_2MnO_4 + 2H_2O$.

44 $C_2H_4 + 3O_2 \longrightarrow 2CO_2 + 2H_2O$.

46 $C + 4HNO_3 \longrightarrow 4NO_2 + CO_2 + 2H_2O$.

48 $5HCl + HClO_3 \longrightarrow 3H_2O + 3Cl_2$.

6.3 Análisis insumo-producto: páginas 734-738

50 *(a)* $\vec{x} = (72, 96)$. *(b)* $\vec{x} = (124, 162)$.

52 *(a)* $\vec{x} = (92, 108)$. *(b)* $\vec{x} = (96, 84)$.

54 *(a)* $\vec{x} = (660, 528, 495)$. *(b)* $\vec{x} = (1\,320, 1\,056, 990)$.

56 *(a)* $\vec{x} = (240, 308, 240)$. *(b)* $\vec{x} = (420, 516, 420)$.

58 136 del sector I_1 y 144 del sector I_2.

60 414 del sector I_1, 483 del sector I_2 y 438 del sector I_3.

62 215 del sector I_1, 270 del sector I_2 y 235 del sector I_3.

64 *(a)* $\vec{x} = (\frac{9}{7}r, \frac{16}{7}r, r)$, $r > 0$. *(b)* $(45, 80, 35)$.

66 $x_1 \approx 0.86x_3$, $x_2 \approx 0.92x_3$.

68 $\nu = 3$,

$$(I_5 - A)^{-1} \approx \begin{bmatrix} 1.0042 & 6.0288\times 10^{-2} & .00189 & 0 & 0 \\ 7.0651\times 10^{-2} & 1.0048 & 3.1719\times 10^{-2} & 0 & 0 \\ 1.2043\times 10^{-2} & 2.1776\times 10^{-2} & 1.0533 & 0 & 0 \\ 0 & 0 & 0 & 1.0533 & 2.1934\times 10^{-2} \\ 0 & 0 & 0 & 3.2901\times 10^{-2} & 1.0423 \end{bmatrix}.$$

6.4 Programación lineal: páginas 738-743

70 Máximo $= 9$ en $(6, 3)$; mínimo $= 4$ en $(3, 1)$.

72 Máximo $= 2\,400$ en $(60, 30)$; mínimo $= 900$ en $(10, 30)$.

74 Mínimo $= 18$ en $(2, 3)$; no alcanza máximo.

76 Máximo $= -8$ en $(3, 2)$; no alcanza mínimo.

78 No alcanza máximo ni mínimo.

80 40 gramos del alimento 1 y 40 gramos del alimento 2 para un costo mínimo de \$10.00.

82 *(a)* $r \leq 1/2$ *(b)* $r \geq 3$.

84 *(a)* 500 kg de mezcla A y 0 Kg mezcla B para una ganancia máxima de \$85 000.
(b) 600 kg de mezcla B y 0 kg de mezcla A para una ganancia máxima de \$48 000.

86 10 kg del fertilizante A y 20 kg del fertilizante B a un costo mínimo de \$350.

88 $P = 24$ en $(0, 3, 0)$.

90 $P = 9$ en $(3, 0, 0)$.

92 $p = 44$ en $(2, 10, 0, 0)$.

94 $P = 70$ en $(0, 5, 15, 0)$.

96 $C = -9$ en $(0, 0, 0, 3)$.

98 200 latas tipo I, 200 latas tipo II y 250 latas tipo III para un ingreso máximo de $38 000.

100 15 del modelo I, 5 del modelo III y 0 del modelo II para una ganancia máxima de $9 000 000.

102 300 unidades de Irapuato y 100 unidades de Monterrey con una utilidad máxima de $290 000.

104 $C = 52$ en $(1, 9, 0)$.

106 $P = 18$ en $(9, 0, 0)$.

109 30 kg, 35 kg y 35 kg de la primera, segunda y tercera sustancias, respectivamente, a un costo mínimo de $1 500.

110 30 unidades de C a B_1, 10 unidades de C a B_2, 0 unidades de D a B_1 y 30 unidades de D a B_2; para un costo de $4 300.

112 3 750 dólares en acciones, 7 500 dólares en bonos y 3 750 dólares en la cuenta de ahorro; para un rendimiento máximo de 1 087.50 dólares.

114 *(a)*

$$\begin{array}{ll} \text{Maximizar} & P = 2x_1 + 6x_2 \\ \text{sujeto a} & \left\{ \begin{array}{rcrcr} 2x_1 & + & 2x_2 & \leq & 3 \\ x_1 & + & x_2 & \leq & 4 \end{array} \right. \\ & x_1, x_2 \geq 0 \end{array}$$

(b) $P = C = 9$, $\vec{x} = (0, 3/2)$, $\vec{y} = (3, 0)$.

116 *(a)*

$$\begin{array}{ll} \text{Minimizar} & C = 8y_1 + 15y_2 \\ \text{sujeto a} & \left\{ \begin{array}{rcrcr} y_1 & + & 3y_2 & \geq & 4 \\ 2y_1 & + & y_2 & \geq & -2 \end{array} \right. \\ & y_1, y_2 \geq 0 \end{array}$$

(b) $P = C = 20$, $\vec{x} = (5, 0)$, $\vec{y} = (0, 4/3)$.

118 *(a)*

$$\begin{array}{ll} \text{Maximizar} & P = 6x_1 + 5x_2 + 8x_3 + 16x_4 \\ \text{sujeto a} & \left\{ \begin{array}{rcrcrcrcr} 2x_1 & + & x_2 & + & x_3 & + & x_4 & \leq & 4 \\ x_1 & + & x_2 & + & 2x_3 & + & 4x_4 & \leq & 3 \end{array} \right. \\ & x_1, x_2, x_3, x_4 \geq 0 \end{array}$$

(b) $P = C$, $\vec{x} = (1, 2, 0, 0)$, $\vec{y} = (1, 4)$.

120 24 días el taller 1 y 38 días el taller 2 con un costo mínimo de $172 000.

6.5 Teoría de juegos: páginas 743-746

122 R juega por filas: *(a)* $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$; *(b)* no es estrictamente determinado.

124 I juega por filas: *(a)* $\begin{bmatrix} 15 & -30 \\ 30 & -20 \end{bmatrix}$ (en miles de dólares); *(b)* estrictamente determinado; *(c)* valor del juego (punto silla) -20000; establecerse ambas en el suburbio.

126 Sí; el proceso I con la materia prima que tiene 4 % del químico.

128 Estrictamente determinado; 2.

130 Estrictamente determinado; 2.

132 Estrictamente determinado; 0.

135 $\begin{bmatrix} 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 8 & 0 \\ -6 & 3 \end{bmatrix} \begin{bmatrix} 1/3 \\ 2/3 \end{bmatrix} = \frac{16}{9}$.

137 $\begin{bmatrix} 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 4 & -2 & 2 \\ -8 & 2 & -4 \end{bmatrix} \begin{bmatrix} 1/4 \\ 1/2 \\ 1/4 \end{bmatrix} = -\frac{3}{4}$.

139 $\vec{p}_0 = (3/4, 1/4)$, $\vec{q}_0 = (5/8, 3/8)$, $\upsilon = 1/4$.

141 $\vec{p}_0 = (1/2, 1/2)$, $\vec{q} = (2/5, 3/5)$, $\upsilon = 1$.

143 $\vec{p}_0 = (1/2, 1/2)$, $\vec{q}_0 = (11/18, 7/18)$, $\upsilon = 3/2$.

145 $\vec{p}_0 = (7/12, 5/12)$, $\vec{q}_0 = (7/12, 5/12)$, para R y T, respectivamente; $\upsilon = 1/12$.

147 $\vec{p}_0 = (1/2, 1/2)$ para los bombarderos y $\vec{q}_0 = (9/28, 17/28)$, donde en ambos vectores la primera componente es la probabilidad de volar alto y la segunda la de volar bajo; $\upsilon = 11/2$.

150 $\vec{q}_0 = (1/9, 0, 8/9)$, $\vec{p}_0 = (2/3, 1/3)$, $\upsilon = 13/3$.

152 $\vec{q}_0 = (0, 5/11, 6/11)$, $\vec{p}_0 = (5/11, 0, 6/11)$, $\upsilon = 14/11$.

154 $\vec{p}_0 = (1/2, 1/2)$, $\vec{q}_0 = (0, 1/3, 2/3)$, $\upsilon = -1$.

156 $\vec{p}_0 = (4/5, 1/5, 0)$, $\vec{q}_0 = (0, 4/5, 1/5)$, $\upsilon = 14/5$.

159 $\vec{p}_0 = (1/3, 1/3, 1/3) = \vec{q}_0$, $\upsilon = 0$.

160 Invertir 0 % en bonos, 30 % en acciones de armamento y 70 % en acciones industriales para un valor esperado de 8.6%.

162 30 bombas de calidad alta y 80 bombas de calidad baja.

6.6 Cadenas de Markov: páginas 747-749

163 *(a)* 43.4 % en *IGN*, 32.6 % en *Guardián del camino* y 24 % en *PGN*.

(b) 43.76 % en *IGN*, 33.12 % en *Guardián del camino* y 23.12 % en *PGN*

166 *(a)* 39.7 % nivel superior, 26.88 % nivel medio y 33.42 % nivel básico.

(b) 53.62 % nivel superior, 24.64 % nivel medio y 21.74 % nivel básico.

168 No existe el límite.

171 *(a)* 20 % ambulatorios, 18.15 % están internados, 42.25 % se han dado de alta y 19.5 % han muerto.

(b) 71.84 % se han recuperado y 28.16 % han muerto.

172 80 % y 20 %, respectivamente.

6.7 Sistemas lineales de ecuaciones diferenciales: páginas 749-756

173 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} C_1e^t + C_2e^{-t} \\ C_1e^t + 2C_2e^{-t} \end{bmatrix}$.

175 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2C_1e^t + 3C_2e^{-t} \\ C_1e^t + C_2e^{-t} \end{bmatrix}$.

177 $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} C_1e^{2t} + C_2e^{-t} \\ C_1e^{2t} + C_3e^{-t} \\ C_1e^{2t} + C_2e^{-t} \end{bmatrix}$.

179 $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} C_1e^t + C_2e^{2t} + C_3e^{3t} \\ C_1e^t + C_2e^{2t} \\ 2C_1e^t + C_2e^{2t} + 2C_3e^{3t} \end{bmatrix}$.

182 $y^2 + x^2 = C^2$.

192 $y = C_1e^x + C_2e^{-3x}$.

194 $y = C_1e^{4x} + C_2xe^{4x}$.

196 $y = C_1e^{2x}\cos 3x + C_2e^{2x}\operatorname{sen} 3x$.

201 $y(x) = \frac{5}{16} + C_1e^{4x} + C_2e^{4x}x$.

202 $y = -5x + C_1 + C_2e^{2x}$.

205 si $\begin{aligned} x_1(t) &= 7 + C_1e^t + C_2e^{-t} \\ x_2(t) &= 10 + C_1e^t + 2C_2e^{-t} \end{aligned}$

207 si $\begin{aligned} x_1(t) &= -5 - C_2e^{-t} + C_1e^t \\ x_2(t) &= -7 - C_2e^{-t} + 2C_1e^t \\ x_3(t) &= 8 + C_2e^{-t} + C_3e^t \end{aligned}$

208 si $\begin{aligned} I_1(t) &= 10 - \tfrac{20}{3}e^{-250t} - \tfrac{10}{3}e^{-1000t} \\ I_2(t) &= \tfrac{10}{3}e^{-250t} - \tfrac{10}{3}e^{-1000t} \\ I_3(t) &= I_1(t) + I_2(t) \end{aligned}$

209 $\begin{aligned} I_1(t) &= \tfrac{1}{2} + \tfrac{5}{6}e^{-80t} - \tfrac{4}{3}e^{-200t} \\ Q(t) &= \tfrac{1}{200} - \tfrac{1}{120}e^{-80t} + \tfrac{1}{300}e^{-200t} \\ I_3(t) &= Q'(t) \\ I_2(t) &= I_1(t) - I_3(t) \end{aligned}$

211 Si x_1 y x_2 son las cantidades instantáneas de sal en sendos tanques A y B,

$$\begin{aligned} x_2(t) &= 600e^{-\frac{1}{40}t} - 600e^{-\frac{3}{40}t} \\ x_1(t) &= 300e^{-\frac{1}{40}t} + 300e^{-\frac{3}{40}t} \end{aligned}$$

213 $\begin{aligned} x_2(t) &= 4000 - 5500e^{-\frac{1}{50}t} + 1700e^{-\frac{3}{50}t} \\ x_1(t) &= 4000 - 2750e^{-\frac{1}{50}t} - 850e^{-\frac{3}{50}t} \end{aligned}$. 4 kilogramos en cada tanque.

215 $x_1(t) = \sqrt{2}C_2e^t \operatorname{sen}\sqrt{2}t - \sqrt{2}C_1e^t \cos\sqrt{2}t$
$x_2(t) = C_1e^t \operatorname{sen}\sqrt{2}t + C_2e^t \cos\sqrt{2}t$

218 $x_1(t) = -\frac{6}{5}\cos\sqrt{6}t + \frac{1}{5}\cos t$
$x_2(t) = \frac{3}{5}\cos\sqrt{6}t + \frac{2}{5}\cos t$

6.8 Optimización de funcionales: páginas 756-760

234 $\{(x,y,z) \mid x^2+y^2+z^2 > 4, x \neq 0, y \neq 0, z \neq 0\}$

236 $\mathbb{R}^3$.

239 $J'(A)H = 2AH$.

241 $J'(A)H = \operatorname{tra}(H)$.

243 $J'(f)h = n(f')^{n-1}h'$.

245 $J'(f)h = 2\langle f,h\rangle$.

246 $df(\vec{u})\vec{h} = 4 - e^{-1}$.

248 $df(\vec{u})\vec{h} = 2$.

250 $df(1,1,1,1)(5,5,5,5) = 8$.

251 $df(\vec{u})\vec{h} = (2,3,0)$.

253 $df(\vec{u})\vec{h} = (0,-1,0,1)$.

255 $df(\vec{u})\vec{h} = k(1,\cos t,-\operatorname{sen} t)$.

258 Mínimo en $(-1/4,-1/4)$.

260 Mínimos relativos en $(-1,-1)$ y $(1,1)$; el criterio no es aplicable en $(0,0)$.

262 Mínimo en $(0,0)$, punto silla en $(-1/4,-1/2)$.

264 Mínimo en $(-2,-3,4)$.

266 Puntos silla en $(0,0,1)$, $(0,0,-1)$ y $(1,1,-1)$; mínimo en $(1,1,1)$.

267 Mínimo en $(1/2,1,1)$. Para encontrar los signos de los valores propios, calcule el polinomio característico de la matriz hessiana y bosqueje su gráfica utilizando derivación.

269 Mínimo en $(1/2,1/4,1/8)$.

270 $f(x) = 3x$.

272 $f(x) = \frac{1}{4}x^2 + \frac{5}{4}x + 1$.

275 $f(x) = e^{x-1}$.

276 $C_1 \cosh\left(\frac{x-C_2}{C_1}\right)$.

279 $f(x) = C_1 \operatorname{sen}(x-C_2)$.

Capítulo 7

7.1 La calculadora HP 50g y álgebra lineal: páginas 813-815

1 *(b)* $-12+2\sqrt{3}$.
(e) $\sqrt{10}-6\sqrt{6}$.

3 *(a)* $\left(\sqrt{3}-\sqrt{2}\sqrt{3}, -\sqrt{3}, -\sqrt{2}+2\right)$.
(c) $-2\sqrt{3}+\sqrt{2}\sqrt{3}-2\sqrt{2}+4$.

5 $a^5 - a + 1$

6 $\begin{bmatrix} a^4-a & -2\cdot a^3 & \frac{a^3+1}{a} \end{bmatrix}$

7 $\{[\ 1\ \ 0\ \ 0\ \ -\frac{7}{22}\ ], [\ 0\ \ 1\ \ 0\ \ -\frac{17}{22}\ ], [\ 0\ \ 0\ \ 1\ \ \frac{19}{11}\ ]\}$.

8 $\begin{bmatrix} 1 & \frac{-83}{67} & \frac{-10}{67} & \frac{-119}{67} & -1 \end{bmatrix}$

10 *(a)* 3. *(b)* $-2a^3$.

11 *(b)* $\begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{2\cdot\sqrt{33}}{33} & -\frac{\sqrt{22}}{22} \\ -\frac{\sqrt{3}}{3} & \frac{4\cdot\sqrt{33}}{33} & \frac{5\cdot\sqrt{22}}{66} \\ \frac{\sqrt{3}}{3} & \frac{2\cdot\sqrt{33}}{33} & \frac{4\cdot\sqrt{22}}{33} \\ 0 & -\frac{\sqrt{33}}{11} & \frac{5\cdot\sqrt{22}}{33} \end{bmatrix}, \begin{bmatrix} \sqrt{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ 0 & \frac{\sqrt{33}}{3} & -\frac{\sqrt{33}}{33} \\ 0 & 0 & \frac{3\cdot\sqrt{22}}{11} \end{bmatrix}$

12 Las columnas de la matriz

$$\begin{bmatrix} \frac{\sqrt{7}}{7} & \frac{9\cdot\sqrt{119}}{119} & -\frac{5\cdot\sqrt{238}}{238} \\ -\frac{\sqrt{7}}{7} & \frac{5\cdot\sqrt{119}}{119} & \frac{\sqrt{238}}{238} \\ \frac{2\cdot\sqrt{7}}{7} & -\frac{3\cdot\sqrt{119}}{119} & -\frac{2\cdot\sqrt{238}}{119} \\ \frac{\sqrt{7}}{7} & \frac{2\cdot\sqrt{119}}{119} & \frac{\sqrt{238}}{17} \end{bmatrix}$$

7.2 MATLAB y álgebra lineal: páginas 815-817

35 $B^t A^2 = \begin{bmatrix} 1986 & -239 & 7146 & 8780 & 9897 \\ -242 & 60 & -1326 & -1869 & -1714 \\ 1039 & -158 & 537 & 860 & 996 \\ 723 & -79 & 129 & 4 & 789 \end{bmatrix}.$

36 $(AB)(AB)^t = \begin{bmatrix} 5354 & -102 & 9399 & 10452 & -6769 \\ -102 & 4884 & -6420 & 11544 & 13636 \\ 9399 & -6420 & 26898 & -21288 & -30656 \\ 10452 & 11544 & -21288 & 303289 & 34517 \\ -6769 & 13636 & -30656 & 34517 & 46867 \end{bmatrix}.$

39 $A^{-1} \approx \begin{bmatrix} -0.1429 & -0.0952 & 0.0952 & 0.0000 & 0.1429 \\ -0.9663 & -0.7844 & 0.9495 & 0.0583 & 0.6588 \\ 0.5261 & 0.0778 & -0.2720 & -0.0097 & -0.0860 \\ -0.2020 & -0.1034 & 0.0743 & 0.0485 & 0.0014 \\ -0.1928 & 0.0333 & 0.1609 & 0.0097 & 0.0860 \end{bmatrix}.$

40 $A^{-1} = \begin{bmatrix} -1/7 & -2/21 & 2/21 & 0 & 1/7 \\ -\frac{2090}{2163} & -\frac{5090}{6489} & \frac{6161}{6489} & \frac{6}{103} & \frac{475}{721} \\ \frac{1138}{2163} & \frac{505}{6489} & -\frac{1765}{6489} & -\frac{1}{103} & -\frac{62}{721} \\ -\frac{437}{2163} & -\frac{671}{6489} & \frac{482}{6489} & \frac{5}{103} & \frac{1}{721} \\ -\frac{139}{721} & \frac{24}{721} & \frac{116}{721} & \frac{1}{103} & \frac{62}{721} \end{bmatrix}.$

42 $\begin{bmatrix} 17.5945 & -4.6920 & -5.5306 \\ 4.9493 & 3.0254 & -0.0544 \\ 64.6095 & -21.0295 & -13.6057 \\ 109.3571 & -32.8816 & -18.4822 \end{bmatrix}$

49 $A^{-1} \approx \begin{bmatrix} -0.0168 & 0.7157 & 0.5837 & -0.3645 \\ -0.0959 & 0.3736 & -0.1933 & 0.1459 \\ 1.0165 & -0.2434 & -0.4226 & 0.2252 \\ -0.2805 & 0.0707 & 0.4643 & 0.0395 \end{bmatrix}.$

50 $A^5 + A = \begin{bmatrix} -45923+7010i & -8096+34676i & -21171+14249i & 54880-30635i \\ 25314-12962i & 16033-47720i & 72842-18000i & 30389+69881i \\ 52069+4020i & 15434+22192i & -73127-52921i & -19562+14867i \\ -49563+31577i & -5254+75206i & -37348+29930i & -1181-33805i \end{bmatrix}$

56 $\det(A) = 2338 + 3744i.$

57 $A^{-1} \approx \begin{bmatrix} 0.1138+0.1728i & -0.0632-0.0314i & 0.0100+0.0097i & -0.0961-0.1369i \\ -0.0251-0.1073i & 0.0669+0.0434i & 0.0646+0.0176i & 0.0587+0.0275i \\ 0.1105-0.0071i & -0.0018+0.0593i & -0.0659+0.0435i & -0.0921+0.0098i \\ 0.0615+0.0816i & 0.0361-0.0304i & -0.0075-0.0415i & 0.0315-0.0172i \end{bmatrix}.$

64 *(a)* $|A| = e$. *(b)* $A^{-1} = \begin{bmatrix} -\frac{f}{e} & 1 & 0 \\ -\frac{g}{e} & 0 & 1 \\ e^{-1} & 0 & 0 \end{bmatrix}$

76 *(a)* Al aplicar a la matriz A la instrucción `rref` se obtiene una matriz con unos en la diagonal principal y ceros fuera de ella, por tanto los vectores columna de la matriz son L.I. y forman una base de la imagen de la transformación; luego el Kernel es el subespacio trivial $\{\vec{0}_{\mathbb{R}^3}\}$.

77 $p = [1.0000 \quad 0.0000 \quad -7.0000 \quad -6.0000]$. $V = \begin{bmatrix} -0.7071 & 0.1961 & 0.3015 \\ 0.0000 & 0.7845 & -0.3015 \\ 0.7071 & 0.5883 & 0.9045 \end{bmatrix}$,

$D = \text{diag}(-1.0000, 3.0000, -2.0000)$.

7.4 Excel, la herramienta Solver y programación lineal: páginas 817-818

95 3.5 en $(1.75, 0, 0)$.

97 115 en $(17.5, 0, 0, 12.5)$.

99 5.5 en $(3, 0, 0, 0.5)$.

101 4 en $(3.2, 0, 0.2, 0)$.

103 15.05263158 en $(0, 3.710526316, 0.131578947, 1.447368421)$.

105 15 en $(0, 0, 0, 1, 7)$.

Capítulo 8

8.1 Aritmética de la computadora y errores de redondeo: página 956

1 *(a)* 3.1415, *(b)* 3.1416.

3 *(a)* 2.718, *(b)* 2.718.

5 *(a)* 0.3826, *(b)* 0.3827.

7 *(a)* 0.000018, *(b)* 0.000018.

9 *(a)* 2 120, *(b)* 0.0058.

11 *(a)* $1/3 \times 10^{-3}$, *(b)* 10^{-3}.

13 *(a)* Menor a 2.136×10^{-4}, *(b)* menor a 1.511×10^{-4}.

15 *(a)* Menor a 7.3465×10^{-6}, *(b)* menor a 2.3385×10^{-6}.

17 *(a)* Menor a 0.00995, *(b)* menor a 0.00366.

19 $(119.99, 120.01)$.

21 $(75.999, 76.001)$.

23 $(5299.9998, 5300.0002)$.

25 $(75.924, 76.076)$.

27 *(a)* $\frac{62}{63}$, *(b)* 0.983, *(c)* 0.985.

29 *(a)* 1, *(b)* 0.999, *(c)* 1.

31 *(a)* $\frac{47}{105}$, *(b)* 0.448, *(c)* 0.447.

8.2 Métodos directos para resolver sistemas lineales: páginas 956-961

33 $\vec{x} = (2, -2, 1, 3)$.

35 El sistema no tiene solución única.

37 $\vec{x} = (1/2, 1/2, -3/2, 1/2, -1/2)$.

40 $A^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 1 & -3 \\ 1 & 0 & -1 \end{bmatrix}$.

42 A es singular.

44 $$A^{-1} = \begin{bmatrix} -\frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 1 & -1 & 0 & -\frac{1}{2} \\ 13 & 9 & -13 & -\frac{3}{2} & -\frac{19}{2} \\ -\frac{5}{2} & -2 & 3 & \frac{1}{2} & 2 \\ -6 & -4 & 6 & \frac{1}{2} & \frac{9}{2} \end{bmatrix}.$$

46 $$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix}, U = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & -8 \end{bmatrix}.$$

48 $$L = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ -2/3 & 1/5 & 1 \end{bmatrix}, U = \begin{bmatrix} 3 & -1 & 4 \\ 0 & 5/3 & 1/3 \\ 0 & 0 & 23/5 \end{bmatrix}.$$

50 $$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 3 & -2 & 1 & 0 \\ 1 & -1 & 1/2 & 1 \end{bmatrix}, U = \begin{bmatrix} 1 & 2 & -4 & 7 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & 10 & -8 \\ 0 & 0 & 0 & -5 \end{bmatrix}.$$

52 $$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 3 & -5 & 1 & 0 & 0 \\ 2 & 2 & -10/31 & 1 & 0 \\ -2 & -1 & -2/31 & -136/33 & 1 \end{bmatrix}, U = \begin{bmatrix} 1 & -1 & 3 & 1 & 2 \\ 0 & -1 & -6 & 1 & -3 \\ 0 & 0 & -31 & 6 & -23 \\ 0 & 0 & 0 & -33/31 & -75/31 \\ 0 & 0 & 0 & 0 & -71/11 \end{bmatrix}.$$

54 $$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 1 & 1 \end{bmatrix}, U = \begin{bmatrix} 1 & -1 & 2 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & -2 \end{bmatrix}.$$

56 $$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 4 & 0 & 1 & 1 \end{bmatrix}, U = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 3 & -3 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

58 $$\vec{x} = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

60 $$\vec{x} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 4 \\ 2 \end{bmatrix}.$$

62 *(a)* $\vec{x} = (12, -0.995)$,
(b) $\vec{x} = (10, -1)$ (solución exacta).

64 *(a)* $\vec{x} = (10, 1)$
(b) $\vec{x} = (10, 1)$ (solución exacta).

66 $\vec{x} = (-1, 2, 4)$.

68 $\vec{x} = (1, -1, 0)$.

70 $\vec{x} = (1, -1, 0, 1)$.

72 $\vec{x} = (-1, 0, 1)$.

74 $\vec{x} = (1, 3, -2)$.

76 $\vec{x} = (-1, -1, -3, -2)$.

78 $\vec{x} = (1, -2, -3, 4)$.

80 $\vec{x} = (1.001, 1.000, 3.001)$.

82 $\vec{x} = (-1, -1, -5)$.

84 $\vec{x} = (-4, 2, 3)$.

86 $\vec{x} = (-1, 2, 0, -1)$.

8.3 Métodos iterativos: páginas 962-967

88 $p = 0, \pm 1$.

90 $p = 1, -\frac{1}{2} \pm \frac{1}{2}\sqrt{5}$.

92 $\vec{p} = (0, 0)$.

94 $\vec{p} = ke^x, k \in \mathbb{R}$.

97 $L = 1$.

107 $\sqrt{n}$

108 $\vec{x}_3 = (-0.999593, 1.000257, -0.999997, -1.000190, -0.999979)$.

110 $\vec{x}_3 = (1.001206, 2.945582, -2.003889, -3.979557, -5.033441)$.

112 $\vec{x}_3 = (0.999991, -0.999996, -2.000023, 3.000008, -1.000004)$.

114 $\vec{x}_4 = (-3.000002, 2.000000, 0.999999, -1.999999, 1.000000, -1.999998)$.

116 $\vec{x}_3 = \begin{bmatrix} -0.999981 \\ 1.999992 \\ 1.000000 \end{bmatrix}$

118 $\vec{x}_5 = \begin{bmatrix} 0.417492 \\ -0.124370 \\ 0.930715 \\ -1.045432 \\ -0.758157 \end{bmatrix}$

120 $\vec{x}_5 = \begin{bmatrix} 1.000942 \\ -0.000966 \\ -0.999709 \\ 0.971441 \\ -0.999418 \end{bmatrix}$

122 $\vec{x}_3 = \begin{bmatrix} 80.000156 \\ 160.000078 \\ -79.999961 \end{bmatrix}$

124 $\vec{x}_3 = \begin{bmatrix} -159.999844 \\ -320.000000 \\ -159.999844 \\ 320.000078 \end{bmatrix}$

126 $\vec{x}_3 = \begin{bmatrix} 79.999863 \\ -80.000029 \\ 80.000098 \\ 79.999941 \end{bmatrix}$

128 $\vec{x}_5 = \begin{bmatrix} 8.000977 \\ -7.999878 \\ 16.000244 \end{bmatrix}$

130 $\vec{x}_4 = \begin{bmatrix} 99.999999 \\ -99.999992 \\ 300.000004 \\ -199.999996 \end{bmatrix}$

132 $\vec{x}_3 = (-1.000003, 0.999999, -1.000000, -1.000000, 1.000000)$.

134 $\vec{x}_3 = (0.999995, 1.999994, -1.000003, -2.000000, -1.000000)$.

136 $\vec{x}_4 = (-1.000001, 3.000000, 4.000000, 2.000000, -1.000000)$.

138 $\vec{x}_3 = (-3.000000, -2.000000, 1.000000, 4.000000, -1.000000, -5.000000)$.

141 $x_3 = (0.999992, -2.000008, -3.000000)$.

143 $\vec{x}_5 = (-1.000494, 2.000137, 3.999922, -3.000027, 2.000040)$.

145 $\vec{x}_4 = (-1.000004, 1.999994, 2.999999, -1.999999, -2.000000)$.

146 $S_3 = \begin{bmatrix} 1.020000 & -0.100000 \\ -0.200000 & 1.020000 \end{bmatrix}$.

148 $S_2 = \begin{bmatrix} 1.000000 & -0.010000 & 0.020000 \\ 0.000000 & 1.000000 & 0.020000 \\ -0.010000 & 0.010000 & 1.000000 \end{bmatrix}$.

150 $\|I_3 - B\| > 1$, posiblemente la matriz B no es invertible.

8.4 Transformaciones de Householder: páginas 968-972

152 $H = \begin{bmatrix} -1/3 & -2/3 & -2/3 \\ -2/3 & 2/3 & -1/3 \\ -2/3 & -1/3 & 2/3 \end{bmatrix}$

154 $H = \begin{bmatrix} -4/9 & \frac{7}{9} & 4/9 \\ \frac{7}{9} & \frac{68}{117} & -\frac{28}{117} \\ 4/9 & -\frac{28}{117} & \frac{101}{117} \end{bmatrix}$

156 $H = \begin{bmatrix} -2/3 & 1/3 & -2/3 \\ 1/3 & \frac{14}{15} & 2/15 \\ -2/3 & 2/15 & \frac{11}{15} \end{bmatrix}$

158 $H = \begin{bmatrix} -3/7 & 2/7 & 6/7 \\ 2/7 & \frac{33}{35} & -\frac{6}{35} \\ 6/7 & -\frac{6}{35} & \frac{17}{35} \end{bmatrix}$

160 $H = \begin{bmatrix} -2/7 & 4/7 & -5/7 & -2/7 \\ 4/7 & \frac{47}{63} & \frac{20}{63} & \frac{8}{63} \\ -5/7 & \frac{20}{63} & \frac{38}{63} & -\frac{10}{63} \\ -2/7 & \frac{8}{63} & -\frac{10}{63} & \frac{59}{63} \end{bmatrix}$

164 $H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -4/5 & 3/5 \\ 0 & 0 & 3/5 & 4/5 \end{bmatrix}$

166 $H=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -2/7 & 3/7 & -6/7 \\ 0 & 3/7 & 6/7 & 2/7 \\ 0 & -6/7 & 2/7 & 3/7 \end{bmatrix}$

168 $H=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -2/3 & 1/3 & -2/3 \\ 0 & 1/3 & \frac{14}{15} & 2/15 \\ 0 & -2/3 & 2/15 & \frac{11}{15} \end{bmatrix}$

170 $H=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1/9 & \frac{8}{9} & -4/9 \\ 0 & 0 & \frac{8}{9} & \frac{13}{45} & \frac{16}{45} \\ 0 & 0 & -4/9 & \frac{16}{45} & \frac{37}{45} \end{bmatrix}$

178 $\det(H)=-1.$

181 *(a)* $\vec{u}=\frac{1}{3\sqrt{14}}(9,-4,5,-2)$ *(b)* $HA=\begin{bmatrix} -7 & -1/7 & \frac{19}{7} & 5/7 \\ 0 & \frac{13}{21} & \frac{22}{63} & -\frac{13}{63} \\ 0 & \frac{31}{21} & \frac{67}{63} & \frac{158}{63} \\ 0 & -\frac{25}{21} & \frac{74}{63} & \frac{25}{63} \end{bmatrix}$

183 $Q=\begin{bmatrix} -1/3 & 2/15 & \frac{14}{15} \\ -2/3 & -\frac{11}{15} & -2/15 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}, R=\begin{bmatrix} -3 & 1 & 3 \\ 0 & -5 & -\frac{23}{5} \\ 0 & 0 & \frac{14}{5} \end{bmatrix}.$

185 $Q=\begin{bmatrix} -2/7 & -\frac{6}{35} & \frac{33}{35} \\ -3/7 & -6/7 & -2/7 \\ 6/7 & -\frac{17}{35} & \frac{6}{35} \end{bmatrix}, R=\begin{bmatrix} -7 & 7 & \frac{29}{7} \\ 0 & -35 & -\frac{109}{35} \\ 0 & 0 & -\frac{13}{35} \end{bmatrix}$

187 $Q=\begin{bmatrix} -1/9 & -\frac{8}{9} & -4/9 \\ 4/9 & -4/9 & \frac{7}{9} \\ -\frac{8}{9} & -1/9 & 4/9 \end{bmatrix}, R=\begin{bmatrix} -9 & 1 & 7/3 \\ 0 & 10 & -1/3 \\ 0 & 0 & 7/3 \end{bmatrix}$

188 $(2,-1,0)$.

190 $(1,1,1)$.

192 $(-1,-2,1)$.

194 $Q=\begin{bmatrix} -1/6 & -\frac{17}{18} & \frac{11}{90} & \frac{23}{90} \\ 1/6 & -\frac{25}{126} & -\frac{599}{630} & -\frac{107}{630} \\ -5/6 & -\frac{1}{126} & -\frac{29}{630} & -\frac{347}{630} \\ -1/2 & \frac{11}{42} & -\frac{59}{210} & \frac{163}{210} \end{bmatrix}, R=\begin{bmatrix} -6 & 7 & 1 \\ 0 & 21 & -1 \\ 0 & 0 & -5 \\ 0 & 0 & 0 \end{bmatrix}$

196 $\vec{x}=(1,-2,0)$

199 $Z=\begin{bmatrix} 13 & -29 & 2 \\ 13 & -10 & 24 \\ 0 & -15 & 36 \end{bmatrix}, H=\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{5}{13} & \frac{12}{13} \\ 0 & \frac{12}{13} & \frac{5}{13} \end{bmatrix}$

201 $Z=\begin{bmatrix} 17 & 31 & 22 \\ 17 & -34 & -17 \\ 0 & 17 & -34 \end{bmatrix}, H=\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{15}{17} & \frac{8}{17} \\ 0 & \frac{8}{17} & \frac{15}{17} \end{bmatrix}$

203 $Z = \begin{bmatrix} -3.0000 & -0.3333 & 3.6932 & 0.4995 \\ -6.0000 & 0.1111 & 1.4934 & 1.0444 \\ 0.0000 & 2.4470 & 0.9528 & 1.4062 \\ 0.0000 & 0.0000 & 2.4062 & -0.0639 \end{bmatrix}$, $H = \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & -0.6667 & 0.3027 & -0.6811 \\ 0.0000 & 0.3333 & -0.6963 & -0.6357 \\ 0.0000 & -0.6667 & -0.6508 & 0.3633 \end{bmatrix}$.

205 $Z = \begin{bmatrix} -3.0000 & -1.2857 & -1.5231 & 0.1796 & 4.2420 \\ -7.0000 & 0.3878 & 1.1504 & 0.5701 & 0.9779 \\ 0.0000 & 4.9252 & -1.1849 & 1.3538 & -3.3074 \\ 0.0000 & 0.0000 & -2.1818 & -0.2062 & -0.0733 \\ 0.0000 & 0.0000 & 0.0000 & 0.1007 & -0.9966 \end{bmatrix}$,

$H = \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & -0.5714 & 0.4221 & 0.5602 & 0.4261 \\ 0.0000 & 0.2857 & -0.6316 & 0.7178 & 0.0652 \\ 0.0000 & -0.7143 & -0.3788 & 0.0044 & -0.5884 \\ 0.0000 & -0.2857 & -0.5286 & -0.4135 & 0.6841 \end{bmatrix}$.

207 $R(\pi/4) = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$, $(0, \sqrt{2})$.

209 $K(\pi/2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $(1, 3)$.

211 $K = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{3}{5} & -\frac{4}{5} & 0 \\ 0 & -\frac{4}{5} & -\frac{3}{5} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $K\vec{x} = \begin{bmatrix} 2 \\ 5 \\ 0 \\ 1 \end{bmatrix}$.

213 $Q = \begin{bmatrix} 3/13 & 4/5 & \frac{36}{65} \\ \frac{4}{13} & -3/5 & \frac{48}{65} \\ \frac{12}{13} & 0 & -\frac{5}{13} \end{bmatrix}$, $R = \begin{bmatrix} 845 & 520 & 4225 \\ 0 & 390 & -8450 \\ 0 & 0 & 4225 \end{bmatrix}$

215 $Q = \begin{bmatrix} \frac{1}{15}\sqrt{15} & -\frac{1}{3}\sqrt{3} & \frac{1}{10}\sqrt{10} & -\frac{1}{2}\sqrt{2} \\ -\frac{1}{15}\sqrt{15} & \frac{1}{3}\sqrt{3} & \frac{7}{30}\sqrt{10} & -\frac{1}{6}\sqrt{2} \\ \frac{2}{15}\sqrt{15} & \frac{1}{3}\sqrt{3} & -\frac{2}{15}\sqrt{10} & -\frac{1}{3}\sqrt{2} \\ \frac{1}{5}\sqrt{15} & 0 & \frac{2}{15}\sqrt{10} & \frac{1}{3}\sqrt{2} \end{bmatrix}$, $R = \begin{bmatrix} \sqrt{15} & \frac{1}{3}\sqrt{15} & -\frac{1}{5}\sqrt{15} \\ 0 & \frac{4}{3}\sqrt{3} & 0 \\ 0 & 0 & \frac{6}{5}\sqrt{10} \\ 0 & 0 & 0 \end{bmatrix}$.

8.5 Aproximación de valores y vectores propios: páginas 972-983

217 $D_1 = \{z : |z-1| \leq 2\}$, $D_2 = \{z : |z+1| \leq 1\}$; $\alpha = 0$, $\beta = 3$

219 $D_1 = \{z : |z-1-3i| \leq 3/4\}$, $D_2 = \{z : |z+2| \leq 5/4\}$, $D_3 = \{z : |z-5| \leq 2\}$; $\alpha = 3/4$, $\beta = 7$

222 $\mu_4 = 4.3241$, $\vec{u}_4 = (1.0000, -0.4990, 0.4907)$

224 $\mu_5 = 5.3338$, $\vec{u}_5 = (1.0000, -0.4764, 0.5237)$.

226 $\mu_3 = -4.5143$, $\vec{u}_3 = (1.0000, -0.4573, 0.5380)$.

235 $\mu_2 = 5.6760$, $\vec{u}_2 = (0.5610, 1.0000, 0.6341)$.

237 $\mu_2 = 6.0244$, $\vec{u}_2 = (1.0000, 0.5139, 0.0556)$.

239 $\mu_2 = 6.2385$, $\vec{u}_2 = (0.6765, 0.6765, 1.0000)$.

242 $\lambda_2 = -2$, $\lambda_3 = -1$.

244 $\lambda_2 = -2$, $\lambda_3 = -1$.

246 $\lambda_2 = -3$, $\lambda_3 = -1$.

248 $\lambda_2 = -3$, $\lambda_3 = 2$, $\lambda_4 = -1$.

250 $\lambda_2 = 2$, $\lambda_3 = -1$, $\lambda_4 = 0$.

252 *(a)* $\mu_2 = -7$, $\vec{u}_2 = (1, 1, 0)$. *(b)* $\lambda_1 = -7$, $\vec{v}_1 = (1, 1, 0)$. *(c)* $\lambda_1 = -7$, $\lambda_2 = -2$, $\lambda_3 = 1$.

254 *(a)* $\mu_3 = -8.6473$, $\vec{u}_3 = (0.9367, 1.0000, -0.0413)$. *(b)* $\lambda_1 = -9$, $\vec{v}_1 = (1,1,0)$. *(c)* $\lambda_1 = -9$, $\lambda_2 = -2$, $\lambda_3 = 1$.

256 *(a)* $\mu_3 = -13.8085$, $\vec{u}_3 = (0.9848, 1.0000, -0.0101)$. *(b)* $\lambda_1 = -14$, $\vec{v}_1 = (1,1,0)$. *(c)* $\lambda_1 = -14$, $\lambda_2 = -3$, $\lambda_3 = 1$.

258 $\lambda_1 = 9$, $\lambda_2 = -3$, $\lambda_3 = 2$, $\lambda_4 = 1$

260 $\lambda_1 = 8$, $\lambda_2 = -4$, $\lambda_3 = -2$, $\lambda_4 = 1$

265 $\vec{u} = (1, 0.8, 0.6)$.

267 $\vec{u} = (0, 1/2, 1)$.

269 $\vec{u} = (1, 1/3, -1/3)$.

271 $\vec{u} = (-2/3, -1/3, 1)$.

273 $\vec{u} = (-2/3, 2/3, 1)$.

275 $\lambda_1 = \rho_2 = 6.2981$, $\vec{x}_1 = \vec{u}_2 = (1, 0.6975, 0.6501)$.

277 $\lambda_1 = \rho_2 = 6.5559$, $\vec{x}_1 = \vec{u}_2(0.7907, 0.3931, 1.0000)$.

279 $\lambda_1 = \rho_2 = 9.8509$, $\vec{x}_1 = \vec{u}_2 = (0.4450, 0.8019, 1.0000)$.

281 $\lambda_1 = \rho_2 = 9.1623$, $\vec{x}_1 = \vec{u}_2 = (1.0000, 0.7208, 0.7208, 1.0000)$.

283 $\lambda_1 = \rho_3 = 9.7072$, $\vec{x}_1 = \vec{u}_2 = (0.3006, 0.2805, 0.2643, 1.0000)$.

286 $\lambda_1 = 7$, $\lambda_2 = 1.5000 + 1.9365i$, $\lambda_3 = 1.5000 - 1.9365i$

288 $\lambda_1 = 11$, $\lambda_2 = i$, $\lambda_3 = -i$, $\lambda_4 = \frac{1}{2} + \frac{\sqrt{3}}{2}i$, $\lambda_5 = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$.

290 $\lambda_1 = 16$, $\lambda_2 = -2 + 3i$, $\lambda_3 = 2 - 3i$, $\lambda_4 = -2.0$.

291 $\lambda_1 = 9$, $\lambda_2 = 2\sqrt{3}$, $\lambda_3 = -2\sqrt{3}$.

292 $\lambda_1 = 7.4290$, $\lambda_2 = -3.1424$, , $\lambda_3 = 1.7134$

294 $\lambda_1 = 11.9213$, $\lambda_2 = 3.4601$, $\lambda_3 = -1.3814$.

296 $\lambda_1 = 6.5971$, $\lambda_2 = -1.3710$, $\lambda_3 = 0.7740$.

298 $\lambda_1 = 4.8678$, $\lambda_2 = 1.3499$, $\lambda_3 = -1.2177$.

300 $\lambda_1 = 7.2372$, $\lambda_2 = -4.2248$, $\lambda_3 = 1.6254$, $\lambda_4 = 0.3622$.

302 $\lambda_3 = 3.6306$, $\lambda_2 = -3.2392$ y $\lambda_1 = 9.6087$.

304 $\lambda_1 = 7.6154$, $\lambda_2 = -1.4942$, $\lambda_3 = 0.8788$

306 $\lambda_1 = 5.4407$, $\lambda_2 = -1.6721$, $\lambda_3 = -9.9779$, $\lambda_4 = 0.2093$.

309 $2, 4, -1$.

311 0.3210, $0.2395 + 1.0904i$, $0.2395 - 1.0904i$.

313 3.1003, 0.1064, 0.2204, 0.5728.

315 $\lambda_1 = 1$, $\vec{v}_1 = \sqrt{2}/2(1,1,0)$; $\lambda_2 = 2$, $\vec{v}_2 = \sqrt{3}/3(-1,1,1)$; $\lambda_3 = 5$, $\vec{v}_3 = \sqrt{6}/6(1,-1,2)$

317 $\lambda_1 = -1.0711$, $\vec{v}_1 = (0.8335, 0.2947, 0.4673)$; $\lambda_2 = 3$, $\vec{v}_2 = (-0.1562, 0.9370, -0.3123)$; $\lambda_3 = 13.0711$, $\vec{v}_3 = (-0.5300, 0.1874, 0.8271)$.

319 $\lambda_1 = -1$ con vector propio correspondiente $\vec{v}_1 = 1/2(1,1,-1,1)$; $\lambda_2 = \lambda_3 = \lambda_4 = 3$, con vectores propios correspondientes $\vec{v}_2 = \sqrt{2}/2(-1,1,0,0)$, $\vec{v}_3 = \sqrt{6}/6(1,1,2,0)$, $\vec{v}_4 = \sqrt{3}/6(-1,-1,1,3)$.

321 $\lambda_1 = -1$, $\vec{v}_1 = (0.6325, 0.3162, -0.3162, 0.6324)$; $\lambda_2 = 4$, $\vec{v}_2 = (-0.3162, 0.6325, -0.6325, -0.3162)$; $\lambda_3 = 5$, $\vec{v}_3 = (0.0000, 0.7071, 0.7071, 0.0000)$; $\lambda_4 = 2$, $\vec{v}_4 = (-0.7071, 0.0000, 0.0000, 0.7071)$.

323 $\lambda_1 = 6$ con vector propio correspondiente $\vec{v}_1 = \sqrt{5}/5(1,1,1,1,1)$; $\lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = 1$ con vectores propios correspondientes $\vec{v}_2 = \sqrt{2}/2(-1,1,0,0,0)$, $\vec{v}_3 = \sqrt{6}/6(-1,-1,2,0,0)$, $\vec{v}_4 = \sqrt{3}/6\,(-1,-1,-1,3,0)$, $\vec{v}_5 = \sqrt{5}/10(-1,-1,-1,-1,4)$.

Lista de símbolos

La siguiente es una lista de abreviaciones y símbolos matemáticos. Los símbolos de uso común se describen sólo brevemente, mientras que los que se definen en el texto —además de una descripción—, incluyen una referencia a la página donde se hace este proceso.

: o $\mid$	en notación de conjuntos, abreviación de la frase *tales que*, p. 985
$\mid$	símbolo para denotar que un número entero (polinomio) divide a otro número entero (polinimio), pp. 995 y 1023
$\in$	símbolo de pertenencia, p. 985
$\notin$	símbolo de no pertenencia, p. 985
$\mathbb{N}$	$= \{1,2,3,\ldots\}$ conjunto de números naturales
$\mathbb{Z}$	$= \{\ldots,-2,-1,0,1,2,\ldots\}$ conjunto de números enteros
$\mathbb{Q}$	$= \{p/q \mid p,q \in \mathbb{Z}, q \neq 0\}$ conjunto de números racionales
$\mathbb{R}$	conjunto de números reales
$\mathbb{C}$	conjunto de números complejos, p. 1011
$\subset$ o $\subseteq$	símbolo de inclusión de conjuntos, p. 986
$\not\subset$ o $\nsubseteq$	símbolo de no inclusión de conjuntos, p. 986
$\subsetneq$ o $\subsetneqq$	símbolo de inclusión propia de conjuntos, p. 987
$A \cup B$	unión de conjuntos, p. 988
$A \cap B$	intersección de conjuntos, p. 988
$\emptyset$	conjunto vacío, p. 989
A^c	complemento de un conjunto, p. 989
$A - B$	diferencia de conjuntos, p. 989
$A \times B$	producto cartesiano, p. 990
$\{A_\lambda\}_{\lambda \in \Lambda}$	familia indexada de conjuntos, p. 992
$\bigcup_{\lambda \in \Lambda} A_\lambda$	reunión de una familia indexada de conjuntos, p. 992
$\bigcap_{\lambda \in \Lambda} A_\lambda$	intersección de una familia indexada de conjuntos, p. 992
$\Rightarrow$	condicional o implicación lógica, p. 994
$\Leftrightarrow$	bicondicional, p. 999
$\forall$	cuantificador lógico: *para todo* o *para cada*
$\therefore$	abreviación de la frase *por lo tanto*
$\sim q$	negación de la afirmación q, p. 997
$f: A \to B$	función f con dominio A y valores en B, p. 133
$x \mapsto f(x)$	asignación a la variable x del valor $f(x)$ bajo la función f
$y' = \frac{dy}{dx}$	primera derivada de la función y
$y'' = \frac{d^2y}{dx^2}$	segunda derivada de la función y
$x \to x_0$	x tiende a x_0
$\lim\limits_{x \to x_0} f(x)$	límite de $f(x)$ cuando x tiende a x_0, p. 680
$\lvert \cdot \rvert$	valor absoluto; en $\mathbb{C}$, p. 1014
$\int f(x)dx$	integral indefinida de la función f
$\int_a^b f(x)dx$	integral definida de la función f
$\sum$	símbolo sumatorio
$\mathfrak{M}_{m \times n}$	matrices reales de tamaño m por n, p. 4
$\mathfrak{M}_{m \times n}(\mathbb{C})$	matrices complejas de tamaño m por n, p. 13
$\mathscr{O}$	matriz cero, p. 7
I_n	matriz identidad, p. 7
$\vec{b}$	matriz columna o vector pp. 8 y 119.
$\operatorname{diag}(\lambda_1,\ldots,\lambda_2)$	matriz diagonal, p. 8
A^t	transposición de una matriz, p. 9
$R_i \leftrightarrow R_j$	intercambio de filas, p. 22
$R_i \leftrightarrow \alpha R_j$	operación de renglón cambio de escala, p. 22

$R_i \leftrightarrow \alpha R_i + \beta R_j$	operación de renglón suma de filas, p. 22
$A\vec{b} = \vec{x}$	sistema lineal con matriz de coeficientes A y término independiente $\vec{b}$, p. 16
$A \sim B$	equivalencia de matrices por filas, p. 23
$[A \mid \vec{b}]$	matriz aumentada, p. 19
A^{-1}	inversa de una matriz, p. 64
$\lvert A \rvert$	determinante de una matriz, p. 77
det	función determinante, pp. 77, 445
$\text{Cof}(A)$	matriz de cofactores de A, p. 83
$\text{Adj}(A)$	matriz adjunta de A, p. 83
(a,b)	pareja ordenada, p. 990
(a,b)	$= \{x \in \mathbb{R} \mid a < x < b\}$, intervalo abierto
$[a,b]$	$= \{x \in \mathbb{R} \mid a \leq x \leq b\}$, intervalo cerrado
$[f,g]$	mínimo común múltiplo, p. 1024
(f,g)	máximo común divisor, p. 1024
$[a,\infty)$	$= \{x \in \mathbb{R} \mid x \geq a\}$
$\mathbb{R}^2$	plano cartesiano, p. 113
$\mathbb{R}^n$	espacio vectorial euclidiano, p. 119
$\lVert \cdot \rVert$	norma de un vector, pp. 115, 120, 248, 303
$\text{d}(\cdot,\cdot)$	distancia entre vectores, pp. 115, 120, 309
$\vec{u} \cdot \vec{v}$	producto punto (escalar o interior) de vectores, pp. 115, 120
$\perp$	símbolo de ortogonalidad (perpendicularidad), pp. 125, 247
$\mathscr{F}(A)$	espacio de funciones reales con dominio A, p. 134
$\vec{0}_{\mathbb{R}^n}$	neutro aditivo de $\mathbb{R}^n$, p. 122
$\vec{0}_{\mathbf{E}}$	neutro aditivo de $\mathbf{E}$, p. 131
$S < \mathbf{E}$	S es *subespacio vectorial de* $\mathbf{E}$, p. 139
$\mathbf{P}$	espacio de polinomios, p. 133
$\mathbb{R}^\infty$	espacio de sucesiones, p. 133
$\mathbf{P}_k$	espacio de polinomios de grado menor o igual a k, p. 142
$C[a,b]$	espacio de funciones continuas, p. 142
$\text{gn}(\vec{u}_1,\ldots,\vec{u}_m)$	espacio generado por los vectores $\vec{u}_1,\ldots,\vec{u}_m$, p. 145
$E_f(A), E_c(A)$	espacios fila y columna, p. 146
$\dim(\mathbf{E})$	dimensión de un espacio, p. 163
$\text{Nul}(A)$	nulidad de una matriz, p. 169
$\text{R}_f(A)$	rango fila, p. 169
$\text{R}_c(A)$	rango columna, p. 169
$\text{Rang}(A)$	rango de una matriz, p. 171
$\vec{u} \times \vec{b}$	producto vectorial, p. 177
$S_1 + S_2$	suma de espacios, p. 186
$S_1 \oplus S_2$	suma directa de espacios, p. 187
$\mathbf{E}/S$	espacio cociente, p. 188
$\mathscr{L}(B)$	espacio generado por un conjunto, p. 190
χ_A	función característica, p. 202
$\mathbf{E} \times \mathbf{F}$	producto de espacios, p. 215
$\langle \cdot,\cdot \rangle$	producto interior, p. 236
$\text{tra}(A)$	traza de una matriz, p. 242
ℓ_2	espacio de sucesiones de cuadrado sumable, p. 243
$\approx$	*aproximadamente*
$\lVert \cdot \rVert_\infty$	norma cúbica/uniforme, pp. 304, 305, 308
$\lVert \cdot \rVert_1$	norma uno, pp. 305, 306
$B(\vec{u}_0, r)$	bola abierta, p. 311
$B[\vec{u}_0, r]$	bola cerrada, p. 311
$S(\vec{u}_0, r)$	esfera, p. 311
$\lVert \cdot \rVert_p$	norma p, pp. 328, 371
$B_n(x)$	polinomio de Bernstein, p. 339
$S^\perp$	complemento ortogonal, p. 389
$C^n[a,b]$	espacio de funciones con derivada hasta el orden n, p. 403
$\mathscr{L}(\mathbf{E},\mathbf{F})$	espacio de transformaciones lineales, p. 419
$\text{Ker}(T)$	núcleo de una transformación, p. 425
$T(S)$	imagen del espacio S bajo la transformación T, p. 427
$\text{Rang}(T)$	rango de una transformación, p. 431
T^{-1}	inversa de una matriz/operador, pp. 64, 433
$[\vec{u}]_{\mathscr{B}}$	vector de coordenadas relativo a la base $\mathscr{B}$, p. 434
$[T]_{\mathscr{B}}$	representación matricial de un operador, p. 442
$\simeq$	relación de similaridad, p. 445
$[T]_{\mathscr{B}_1}^{\mathscr{B}_2}$	representación matricial de una transformación, p. 447
$\cong$	relación de isomorfismo, p. 452
$C^\infty(-\infty,\infty)$	espacio de funciones con derivada de todo orden, p. 459
E_λ	espacio propio, p. 461
$p_T(\lambda)$	polinomio característico, pp. 467, 468
$\mu_a(\lambda)$	multiplicidad algebraica, p. 475
$\mu_g(\lambda)$	multiplicidad geométrica, p. 475
$\overline{z}$	conjugado de un número complejo, pp. 492 y 1014
$T \circ S$	operador composición, p. 504
TS	$= T \circ S$, composición de transformaciones, pp. 504, 559
I	operador identidad, p. 505
$T^{-1}(S)$	imagen inversa, p. 501
T^n	composición de T consigo mismo n veces, pp. 510, 511
E^*	espacio dual, p. 516

δ_{ij}	delta de Kronecker, p. 517
T^*	operador adjunto, p. 523
S^0	espacio anulador, p. 524
$\text{grad}(p)$	grado de un polinomio, pp. 526, 1022
$\binom{n}{k}$	$= \dfrac{n!}{k!(n-k)!}$, coeficiente binomial
$\mathfrak{m}(\lambda)$	polinomio mínimo, p. 536
T^t	transpuesta de una transformación, p. 565
$\overline{M}$	conjugado de una matriz, p. 576
$P_i \to P_l \cdots \to P_j$	k-cadena de un digrafo, p. 582
$\mathscr{A}(B)$	matriz asociada, p. 585
—|⊢—	símbolo para representar una batería, p. 591
⊂⊢⊃	dirección hacia donde fluye una corriente eléctrica de una batería, p. 591
—/\/\/\/—	resistor eléctrico, p. 591
$\longrightarrow$	símbolo *produce*, en ecuaciones químicas, p. 595
$[\vec{u},\vec{v}]$	segmento entre dos vectores, p. 616
máx	función máximo, p. 630
mín	función mínimo, p. 630
$E(\vec{p},\vec{q})$	valor esperado de un juego, pp. 646, 647
$\dot{x}$	$= dx/dt$, derivada de una función, p. 667
$\dot{\vec{x}} = A\vec{x}$	sistema lineal de ecuaciones diferenciales, p. 667
$\Vert\cdot\Vert_n$	norma en $C^n[a,b]$, p. 681
π_k	k-ésima proyección, p. 684
$\text{ínf}(A)$	ínfimo de un conjunto, p. 687
$dJ(f_0)$ o $J'(f_0)$	diferencial de una función J, p. 689
$\partial_u J(f_0)$	derivada direccional, p. 693
$\dfrac{\partial f}{\partial x_i}$ o f_{x_i}	derivada parcial, p. 698
$[df(\vec{a})]$	matriz jacobiana, pp. 699, 705
$f_{x_i x_j}$	segundas derivadas parciales, p. 709
$H(\vec{a})$	matriz hessiana, p. 710
$\Vert\cdot\Vert^*$	normas matriciales naturales, p. 855
$\rho(A)$	radio espectral, p. 878
$\text{sign}(a)$	función *signo*, p. 904
$\text{máxc}(\vec{x})$	primera componente de $\vec{x}$ con mayor valor absoluto, p. 926
$\text{Re}(z)$, $\text{Im}(z)$	partes real e imaginaria de un número complejo, p. 1013
$\arg(z)$	conjunto de argumentos de un número complejo, p. 1015
$\text{Arg}(z)$	argumento principal de un número complejo, p. 1015
$e^{i\theta}$	$= \cos\theta + i\,\text{sen}\,\theta$, fórmula de Euler, p. 1017
$\mathbb{Z}_2$	campo de números binarios, p. 1018
$\equiv \bmod n$	congruencia módulo n, p. 1018
$[a]$	clase de equivalencia módulo n, p. 1018
$\mathbb{Z}_n$	conjunto de clases de equivalencia módulo n, p. 118
$0_{\mathbb{K}}$, $1_{\mathbb{K}}$	neutros aditivo y multiplicativo de un campo, p. 1017
a^{-1}	inverso multiplicativo de un campo, p. 1017
$\mathbb{K}[x]$	conjunto de polinomios sobre un campo $\mathbb{K}$, p. 1022
$\mathbb{K}^n$	espacio de n-eadas ordenadas con componentes en $\mathbb{K}$, p. 1027
$\text{card}(A)$	cardinalidad de un conjunto, p. 1027
$\aleph_0$	aleph cero, cardinalidad de $\mathbb{N}$, p. 1028
$\mathfrak{c}$	cardinalidad del continuo, p. 1028
$\mathbb{Z}_2^n$	espacio de vectores de código binario, pp. 1029, 1030
$(\mathbf{n},\mathbf{k})$	código (n,k) de Hamming, p. 1034
$\preceq$	relación *precede*, p. 1046
$J_k(\mu)$	bloque de Jordan, p. 1055
N_k	matriz nilpotente básica, p. 1055

Alfabeto griego

May.	Min.	Nombre	May.	Min.	Nombre	May.	Min.	Nombre
A	α	Alfa	I	ι	Iota	P	ρ	Rho o ro
B	β	Beta	K	κ	Kappa	Σ	σ	Sigma
Γ	γ	Gamma	Λ	λ	Lambda	T	τ	Tau
Δ	δ	Delta	M	μ	My o mu	Υ	υ	Ípsilon
E	ε	Épsilon	N	ν	Ny o nu	Φ	ϕ, φ	Phi o fi
Z	ζ	Zeta	Ξ	ξ	Xi	X	χ	Ji
H	η	Eta	O	o	Omicron	Ψ	ψ	Psi
Θ	θ	Theta	Π	π	Pi	Ω	ω	Omega

Lista de aplicaciones adicionales

Esta lista contiene un conjunto de aplicaciones, adicionales al capítulo 6, que se encuentran integradas en algunos capítulos del texto.

Lista de programas

El siguiente es un listado de los principales programas en MATLAB que contiene el texto.

Bibliografía

[1] Aleksandrov, A., Kolmogorov, A., Laurentiev, M. y otros, *La matemática: su contenido, métodos y significado*, volúmenes 2 y 3, Alianza Universidad, Madrid, 1982.

[2] Apostol, T., *Calculus*, volumen 2, Editorial Reverté, S.A., España, 1980.

[3] Burden, R. y Faires, J., *Numerical analysis*, Brooks/Cole, Kentucky, 2001.

[4] Cartan, H., *Cálculo diferencial*, Ediciones Omega, Barcelona, 1978.

[5] Cheney, W. y Kincaid, D., *Análisis numérico: las matemáticas del cálculo científico*, Addison-Wesley Iberoamericana, México, 1994.

[6] Ewing, G., *Calculus of variations with applications*, Dover Publications, Inc., Nueva York, 1985.

[7] Fraleigh, J. y Beauregard, R., *Linear algebra*, Addison-Wesley Pub. Co., Reading, Massachussets, 1990.

[8] Friedberg, S., Insel, A. y Spence, L., *Linear algebra*, Prentice-Hall, Nueva Jersey, 2003.

[9] Halmos, P., *Naive set theory*, Springer-Verlag, Nueva York, 1974.

[10] Herstein, I. N., *Topics in algebra*, John Wiley & Sons, Nueva York, 1975.

[11] Hoffman, K. y Kunze, R., *Álgebra lineal*, Prentice-Hall Internacional, España, 1982.

[12] Ize, J., *Cálculo de variaciones*, IIMAS-FENOMEC, UNAM, México, 2002.

[13] Jones, G., y Jones, J. M., *Information and coding theory*, Springer-Verlag, Nueva York, 2000.

[14] Kreyszyg, E., *Introductory functional analysis with applications*, John Wiley & Sons, Inc., Nueva York, 1978.

[15] Lang, S., *Linear algebra*, Addison-Wesley Pub. Co., Inc., Nueva York, 1971.

[16] Lang, S., *Introducción al análisis matemático*, Addisson-Wesley Iberoamericana, México, 1990.

[17] Leon, S., *Linear algebra with applications*, Macmillan College Pub. Co., Inc., Nueva York, 1994.

[18] Lipschutz, S., *Theory and problems of set theory and related topics*, Schaum Publishing Co, Nueva York, 1964.

[19] Lipschutz, S., *Álgebra lineal*, McGraw-Hill Interamericana, México, 1991.

[20] Pless, V., *Introduction to the theory of error-correcting codes*, John Wiley & Sons, Inc., Nueva York, 1982.

[21] Rivlin, T., *An introduction to the approximation of functions*, Dover Publications, Inc., Nueva York, 1981.

[22] Rojo, J. y Martin, I., *Ejercicios y problemas de álgebra lineal*, McGraw-Hill, Madrid, 1994.

[23] Sundarapandian, V., *Numerical linear algebra*, Prentice-Hall of India, Nueva Deli, 2008.

[24] Wilkinson, J. H., *The algebraic eigenvalue problem*, Oxford University Press, Oxford, 1965.

Índice analítico

Los números de página seguidos de una “n” indican que la entrada se encuentra en las notas.

D

E

F

G

N

O

P

R

S

W

Z